

AAAG67134
ID AAG67134 standard; Protein; 369 AA.
XX AC AAG67134;
XX DT 13-NOV-2001 (first entry)
XX DE Amino acid sequence of a human enzyme.
XX Human; enzyme; cancer; neurological disorder; epilepsy; stroke;
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
KW multiple sclerosis; Parkinson's disease; amyotrophic lateral sclerosis;
KW meningitis; schizophrenia; neuroskeletal disorder; allergy;
KW Addison's disease; autoimmune disease; anemia; asthma; Crohn's disease;
KW adult respiratory distress syndrome; atopic dermatitis; psoriasis;
KW diabetes mellitus; osteoporosis; pancreatitis; rheumatoid arthritis;
KW Huntington's chorea; muscular dystrophy; Gaucher's disease;
KW Von Willebrand's disease; sickle cell anemia; thalassemia; atherosclerosis;
KW leukemia; hepatitis; cirrhosis; arteriosclerosis; gene therapy.
XX OS Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 200
FT /note= "potential phosphorylation site"
FT Modified-site 222
FT /note= "potential phosphorylation site"
FT Modified-site 224
FT /note= "potential phosphorylation site"
FT Modified-site 245
FT /note= "potential phosphorylation site"
FT Modified-site 283
FT /note= "potential phosphorylation site"
FT Modified-site 293
FT /note= "potential phosphorylation site"
FT Modified-site 311
FT /note= "potential phosphorylation site"
FT Modified-site 318
FT /note= "potential phosphorylation site"
FT Modified-site 366
FT /note= "potential phosphorylation site"
XX W0200164896-A2.
XX 07-SEP-2001.
XX 01-MAR-2001; 2001WO-US06806.
XX 01-MAR-2000; 2000US-0186307.
XX 28-MAR-2000; 2000US-0192532.
XX 30-MAR-2000; 2000US-0193578.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Lu DAM, Bandman O, Yue H, Azimzai Y, Lal P, Burford N;
XX Baughn XR;
XX WPI; 2001-550184/61.
XX N-PSDB; AAH75162.
XX Novel human enzyme molecule useful for treating and preventing, e.g.,
XX cancer, genetic disorders, neurological disorders, autoimmune and
XX inflammatory disorders -
XX Claim 1; Page 123-124; 154pp; English.
XX The present sequence represents a human enzyme. The enzyme polynucleotide
XX and polypeptide are useful for diagnosis, treatment and prevention of
XX cancers, neurological disorders (e.g. epilepsy, stroke, Alzheimer's
XX disease, Pick's disease, Huntington's disease, dementia, multiple
XX sclerosis, Parkinson's disease, amyotrophic lateral sclerosis, bacterial
XX and viral meningitis, schizophrenic disorders and neuroskeletal

RESULT 6
AAAY49421
ID AAAY49421 standard; Protein; 304 AA.
XX AC AAAY49421;
XX DT 13-MAR-2000 (first entry)
XX DE Human TREX1h (Exo 1) polypeptide.
XX 3'-5' exonuclease; gene therapy; protocols. antineoplastic; antiviral;
KW chemotherapeutic; cell killing; human; TREX1h; Exo 1.
XX OS Homo sapiens.
XX WO9961064-A1.
XX PD 02-DEC-1999.
XX 14-MAY-1999; 98WO-US10578.
XX 22-MAY-1998; 98US-0083617.
XX 12-NOV-1998; 98US-0191470.
XX (UTWA-) UNIV WAKE FOREST.
XX Perrino FW;
XX WPI; 2000-097077/08.
XX N-PSDB; AAZ46492.
XX Novel genes used in hybridization assays to detect the capacity of
XX cells to express exonucleases -
XX Claim 1; Page 45-47; 93pp; English.
XX The invention relates to human and mouse genes encoding 3'-5'
XX exonucleases. The exonuclease proteins of the invention can be used to
XX identify inhibitors and effectors of exonuclease activity. Specific
XX binding polypeptides, e.g. antibodies, can be used for purifying
XX exonuclease products and detection and quantification of exonuclease
XX products in fluid and tissue samples using immunological procedures.
XX Binding proteins are also useful in modulating the activity of
XX exonucleases. Polynucleotides of the invention are useful in
XX hybridization assays to detect the capacity of cells to express
XX exonucleases. They are also useful as the basis for diagnostic methods
XX useful for identifying a genetic alteration in an exonuclease locus that
XX underlies a disease state. Nucleic acids that modulate the expression of
XX the exonuclease genes, e.g. antisense nucleic acids, ribozymes, triple
XX helix oligonucleotides, can be used in gene therapy protocols. Resistance
XX or ineffectiveness of certain antineoplastic and antiviral agents may be
XX due to an exonuclease activity. A need exists for the identification of
XX metabolic factors which modulate the ability of chemotherapeutic agents
XX to effect cell killing. This need is met by the present invention. The
XX exonuclease polypeptides of the invention can be used to design and
XX identify therapeutics which increase efficiency of the chemotherapeutic
XX agent at lower doses, which are more easily tolerated in patients, and
XX reducing side effects. The present sequence represents a human TREX1h
XX (Exo 1) polypeptide.
XX Sequence 304 AA;
XX Query Match 80.5%; Score 33; DB 21; Length 304;
XX Best Local Similarity 71.4%; Pred. NO. 1.2e-02;
XX Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 TKHGPRK 7
XX Db 159 SERGPRK 165
XX RESULT 7

QY	2	KHGPRK	7
Db	292	KHGPRK	297
RESULT 9			
	ABG15782		
	ID	ABG15782 standard; Protein; 526 AA.	
XX	AC	ABG15782;	
XX	AC		
XX	DT	18-FEB-2002 (first entry)	
XX	DT		
XX	DE	Novel human diagnostic protein #15773.	
XX	DE		
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder.	
XX	XS		
XX	OS	Homo sapiens.	
XX	XX		
XX	PN	WO200175067-A2.	
XX	PD		
XX	PD	11-OCT-2001.	
XX	XX		
XX	PF	30-MAR-2001; 2001WO-US08631.	
XX	PR		
XX	PR	31-MAR-2000; 2000US-0540217.	
XX	PR	23-AUG-2000; 2000US-0649167.	
XX	XX		
XX	PA	(HYSE-) HYSEQ INC.	
XX	XX		
XX	PI	Dmanac RT, Liu C, Tang YT;	
XX	PI		
XX	DR	WPI; 2001-639362/73.	
XX	DR	N-PSDB; AAS79969.	
XX	XX		
PT	XX	New isolated polynucleotide and encoded polypeptides, useful in	
PT	XX	diagnostics, forensics, gene mapping, identification of mutations	
PT	XX	responsible for genetic disorders or other traits and to assess	
PT	XX	biodiversity -	
XX	PS		
XX	PS	Claim 20; SEQ ID NO 46141; 103pp; English.	
XX	XX		
CC	XX	The invention relates to isolated polynucleotide (I) and	
CC	XX	polypeptide (II) sequences. (I) is useful as hybridisation probes,	
CC	XX	polymerase chain reaction (PCR) primers, oligomers, and for chromosome	
CC	XX	and gene mapping, and in recombinant production of (II). The	
CC	XX	polynucleotides are also used in diagnostics as expressed sequence tags	
CC	XX	for identifying expressed genes. (I) is useful in gene therapy techniques	
CC	XX	to restore normal activity of (II) or to treat disease states involving	
CC	XX	(II). (II) is useful for generating antibodies against it, detecting or	
CC	XX	quantitating a polypeptide in tissue, as molecular weight markers and as	
CC	XX	a food supplement. (II) and its binding partners are useful in medical	
CC	XX	imaging of sites expressing (II). (I) and (II) are useful for treating	
CC	XX	disorders involving aberrant protein expression or biological activity.	
CC	XX	The polypeptide and polynucleotide sequences have applications in	
CC	XX	diagnostics, forensics, gene mapping, identification of mutations	
CC	XX	responsible for genetic disorders or other traits to assess biodiversity	
CC	XX	and to produce other types of data and products dependent on DNA and	
CC	XX	amino acid sequences. ABG00010-ABG30377 represent novel human	
CC	XX	diagnostic amino acid sequences of the invention.	
CC	XX	Note: The sequence data for this patent did not appear in the printed	
CC	XX	specification, but was obtained in electronic format directly from WIPO	
CC	XX	at ftp.wipo.int/pub/published_pct_sequences.	
XX	XX		
SQ		Sequence 526 AA;	
Query Match 80.5%; Score 33; DB 22; Length 526;			
Best Local Similarity 85.7%; Pred. No. 2.2e+02;			
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0			
QY	1	TKHGPRK	7

Db 296 TKHTPRK 302

RESULT 10
ABG17605
ID ABG17605 standard; Protein; 526 AA.
XX AC ABG17605;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #17596.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HVSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX DR N-PSDB; AAS81792.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX PS Claim 20; SEQ ID No 47964; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 526 AA;
Query Match 80.5%; Score 33; DB 22; Length 526;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TKHGPRK 7
Db 296 TKHTPRK 302

RESULT 11
ABJ18912
ID ABJ18912 standard; Protein; 717 AA.
XX AC ABJ18912;
XX DT 06-MAR-2003 (first entry)
XX DE Pathogen specific antigen related staphylococcal protein SEQ ID No 58.
XX KW Antibacterial; virucide; fungicide; protozoicide; cytostatic; anti-HIV;
XX KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX KW autoimmune disease; HIV; hepatitis.
XX OS Staphylococcus sp.
XX PN WO200259148-A2.
XX PD 01-AUG-2002.
XX PF 21-JAN-2002; 2002WO-EP00546.
XX PR 26-JAN-2001; 2001AT-0000130.
XX PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX PI Meinke A, Nagy E, Von Absen U, Klade C, Henics T, Zauner W;
XX PI Minh DB, Wytvtska O, Etz H, Dryia A, Weichhart T, Hafner M;
XX PI Tempelmaier B;
XX DR WPI; 2003-075410/07.
XX PT Identifying, isolating and producing hyperimmune serum-reactive
PT antigens from a pathogen, for preparing vaccine or medicament for
PT treating or preventing e.g. staphylococcal infections, comprises
PT providing antibody preparation -
XX PS Claim 24; Page 155; 252pp; English.

CC The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens
CC are used in a vaccine, comprises providing antibody preparation from a
CC plasma pool of a type of animal, or individual sera with antibodies
CC against the specific pathogen, tumour, allergen, tissue or host prone to
CC auto-immunity. The hyperimmune serum-reactive antigens comprising any of
CC the 62 sequences of 53-2261 amino acids fully defined in the
CC specification, or their hyperimmune fragments are useful for the
CC manufacture of a pharmaceutical preparation, particularly a vaccine
CC against staphylococcal infections or colonisation against S. aureus or S.
CC epidermidis. The preparation of antibodies is useful for the manufacture
CC of a medicament for treating or preventing staphylococcal infections or
CC colonisation against S. aureus or S. epidermidis. The antibody
CC preparations may also be used for diagnostic and imaging purposes. Other
CC conditions that can be treated include cancer, autoimmune diseases or
CC infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or
CC protozoan pathogens. This sequence represents a staphylococcal protein
CC relating to the method for identifying and producing pathogen specific
CC antigens of the invention.

XX SQ Sequence 717 AA;
Query Match 80.5%; Score 33; DB 24; Length 717;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPR 6
Db 91 TKHGPK 96

Thu Feb 12 05:07:43 2004

AS 03 JUL 2010 01:43

```

RESULT 12
ABUS2986
ID ABUS2986 standard; Protein; 743 AA.
XX AC
XX ABUS2986;
XX DT 14-APR-2003 (first entry)
XX DE Human nucleic acid management-associated protein from DKF2phtes3_15j3.
XX KW Human; gene therapy; vaccine; disease treatment; detection.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-IB01496.
XX PR 18-AUG-1999; 99US-0149499.
XX PR 28-SEP-1999; 99US-0156503.
XX PR (GSHU-) GERMAN HUMAN GENOME PROJECT.
XX PA
XX PI Wiemann S;
XX DR WPI; 2001-327840/34.
XX DR N-PSDB; ABX71347.
XX PT Nucleic acids having the sequences of clones isolated from libraries of
XX PT different human tissues, useful in recombinant DNA methodologies -
XX PS Claim 21; Page 602; 1095pp; English.
XX CC This invention describes novel polynucleotides and polypeptides isolated
XX CC from human cDNA libraries which can be used for gene therapy or in
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. The products of the
XX CC invention may also be used to identify modulators of expression and
XX CC activity and to down regulate expression and activity. The antibodies of
XX CC the invention may also be used as diagnostic agents for detecting the
XX CC presence of polypeptides in samples. This sequence represents a
XX CC polypeptide described in the disclosure of the invention.
XX SQ Sequence 743 AA;
XX Query Match 80.5%; Score 33; DB 22; Length 743;
XX Best Local Similarity 83.3%; Pred. No. 3.1e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 KHGPRK 7
XX DB 381 KHGPRK 386
XX RESULT 13
AAG80778
ID AAG80778 standard; Protein; 774 AA.
XX AC
XX AAG80778;
XX DT 19-APR-2002 (first entry)
XX DE Human exonuclease factor NEF protein.
XX KW Human; exonuclease factor; NEF; testicular; immunisation;
XX KW antibody preparation.
XX OS Homo sapiens.
XX PN
XX PD
XX PF
XX PR
XX PR 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX PR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55035.

```

```

PN CNI316433-A.
XX 10-OCT-2001.
XX PF 11-APR-2001; 2001CN-0108200.
XX PR 11-APR-2001; 2001CN-0108200.
XX PA (UYNA-) UNIV NANJING MEDICAL.
XX PI Sha J, Zhou Z, Li J;
XX WPI; 2002-067605/10.
XX DR N-PSDB; ABA97189.
XX PT Human exonuclease factor gene encoding a protein -
XX Claim 1; Page 2 (Disclosure); 6pp; Chinese.
XX CC This invention describes a novel human testicular exonuclease factor gene
XX CC (NEF). The NEF gene described in the disclosure if the invention can be
XX CC used to prepare a fusion protein which can be used to immunise animals to
XX CC prepare monoclonal and polyclonal antibodies. The protein of the
XX CC invention can be used to prepare medicines. This sequence represents the
XX CC human NEF protein described in the invention.
XX SQ Sequence 774 AA;
XX Query Match 80.5%; Score 33; DB 23; Length 774;
XX Best Local Similarity 83.3%; Pred. No. 3.2e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 KHGPRK 7
XX DB 381 KHGPRK 386
XX RESULT 14
AAU37176
ID AAU37176 standard; Protein; 897 AA.
XX AC AAU37176;
XX DT 14-FEB-2002 (first entry)
XX DE Staphylococcus aureus cellular proliferation protein #1346.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX PR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55035.

```


XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12769; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 897 AA;

Query Match 80.5%; Score 33; DB 22; Length 897;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPR 6
|||||
DB 307 TKHGPK 312

RESULT 15
AAU34340
ID AAU34340 standard; Protein; 1113 AA.

XX AAU34340;

XX 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #616.

XX Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS52199.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 5836; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1113 AA;

Query Match 80.5%; Score 33; DB 22; Length 1113;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPR 6
|||||
DB 523 TKHGPK 528

Search completed: February 11, 2004, 17:02:50
Job time : 34.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL.23.*

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp mhc.*
- 8: sp organelle.*
- 9: sp phase.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp rvirus.*
- 16: sp bacterioph.*
- 17: sp archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	116	3	013550 saccharomyc
2	34	100.0	168	11	Q8BQA6
3	34	100.0	186	12	Q8JUE6
4	34	100.0	186	12	Q91SF9
5	32	94.1	339	16	Q31702
6	32	94.1	478	16	Q9CJ31
7	32	94.1	905	10	Q8RYR1
8	32	94.1	1381	16	Q8YWH7
9	31	91.2	184	12	Q91R55
10	31	91.2	217	12	Q9YPO9
11	31	91.2	217	12	Q9VPR0
12	31	91.2	263	16	Q8G5N0
13	31	91.2	348	16	Q9CPI0
14	31	91.2	509	17	Q96XL2
15	31	91.2	669	16	Q66983
16	31	91.2	780	5	Q81122

17	31	91.2	822	11	Q8BYW3	Q8byw3 mus musculu
18	31	91.2	2999	11	Q8CHI7	Q8chi7 mus musculu
19	31	91.2	3035	11	Q8CHI8	Q8chi8 mus musculu
20	30	88.2	151	16	Q9KAF5	Q9kaf5 bacillus ha
21	30	88.2	153	16	Q8EQL6	Q8eq16 oceanobacil
22	30	88.2	161	5	Q8WQ09	Q8wq09 oestertagia
23	30	88.2	202	8	Q9TIP1	Q9tip1 pedicularis
24	30	88.2	247	2	O30393	O30393 neisseria g
25	30	88.2	249	16	Q8EXP5	Q8exp5 leptospira
26	30	88.2	255	5	Q8WQ10	Q8wq10 oestertagia
27	30	88.2	332	17	Q9UX18	Q9ux18 sulfolobus
28	30	88.2	396	16	Q926X4	Q926x4 listeria in
29	30	88.2	396	16	Q8Y306	Q8y3g6 listeria mo
30	30	88.2	447	2	Q8GDN0	Q8gdno photorhabd
31	30	88.2	453	16	Q8D319	Q8d319 wiggleswort
32	30	88.2	620	5	Q8MXG4	Q8mxg4 caenorhabdi
33	30	88.2	723	16	Q8PH66	Q8ph66 xanthomonas
34	30	88.2	723	16	Q8F5T8	Q8fst8 xanthomonas
35	30	88.2	735	16	Q9PGF0	Q9pgf0 xyella fas
36	30	88.2	809	10	Q8VYCI	Q8vycl arabidopsis
37	30	88.2	836	10	O24527	O24527 arabidopsis
38	30	88.2	966	5	Q8T9K4	Q8t9k4 drosophila
39	30	88.2	966	5	Q9VLM8	Q9v1m8 drosophila
40	30	88.2	966	5	Q9U6B4	Q9u6b4 drosophila
41	30	88.2	1120	10	Q9LQAI	Q9lqal arabidopsis
42	30	88.2	1222	10	Q8H6V1	Q8h6v1 zea mays (m
43	30	88.2	1222	10	Q8H6U6	Q8h6u6 zea mays (m
44	30	88.2	1226	10	Q8H6V0	Q8h6v0 zea mays (m
45	30	88.2	1247	10	Q8H6U9	Q8h6u9 zea mays (m

ALIGNMENTS

RESULT 1
013550 PRELIMINARY; PRT; (116 AA.)
AC 013550;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE YLR269CP.
GN YLR269C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Kotter P.,
RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Furnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
PT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:0-0(0).

SEQUENCE FROM N.A.
STRAIN=S288C;
Miller N.;
Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
STRAIN=S288C;
Waterston R.;

RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RF SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Cherry J.M.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17244; AAB67385.1; -
 DR SGD: S0004259; YLR269C.
 SQ SEQUENCE 116 AA; 13718 MW; 94F7025D903552E8 CRC64;

Query Match 100.0%; Score 34; DB 3; Length 116;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 67 SLKRLPK 73

RESULT 2
 ID Q8BOA6 PRELIMINARY; PRT; 168 AA.

AC Q8BOA6; 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
 RX MEDLINE=22334693; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL: AK051146; BAC34536.1; -
 DR Hypothetical protein.
 KW Hypothetical protein
 SQ SEQUENCE 168 AA; 18003 MW; AF9BB01F4359A158 CRC64;

Query Match 100.0%; Score 34; DB 11; Length 168;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 120 SLKRLPK 126

RESULT 3

ID Q8JUJ6 PRELIMINARY; PRT; 186 AA.
 AC Q8JUJ6; 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Matrix protein M2.

GN M2.
 OS Turkey rhinotracheitis virus (TRTV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
 CX NCBI_TaxID=11264;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=UK/11/94;
 RA Jacobs J.A., Njenga K., Mawditt K., Britton P., Cavanagh D., Seal B.;
 RT "Nucleotide and predicted amino acid sequence analysis of the
 RT phosphoprotein, second matrix and small hydrophobic protein genes for
 RT avian metapneumovirus type B";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ492378; CAD37362.1; -
 DR InterPro: IPR000571; Znf.CCCH.
 DR Pfam: PF00642; zf-CCCH; I.
 SQ SEQUENCE 186 AA; 20873 MW; 28D549FCA5ED605D CRC64;

Query Match 100.0%; Score 34; DB 12; Length 186;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 146 SLKRLPK 152

RESULT 4

ID Q91SF9 PRELIMINARY; PRT; 186 AA.

AC Q91SF9; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Matrix glycoprotein M2.

GN M2.
 OS Avian pneumovirus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
 CX NCBI_TaxID=38525;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Hungary/657/4;
 RA Seal B.S., Jacobs J.A., Turpin E.A., Njenga M.K.;
 RT "Sequence comparison of avian pneumovirus isolates from the United
 RT States confirms a subgroup different strains."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF356650; AAK48888.1; -
 DR InterPro: IPR000571; Znf.CCCH.
 DR Pfam: PF00642; zf-CCCH; I.
 SQ SEQUENCE 186 AA; 20802 MW; 6214C6187B7711B8 CRC64;

Query Match 100.0%; Score 34; DB 12; Length 186;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 146 SLKRLPK 152

RESULT 5

ID Q31702 PRELIMINARY; PRT; 339 AA.

AC Q31702; 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Molybdopterin biosynthesis protein.

GN MOEB
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunz F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.U., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Estian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.P., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kaahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapins A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nock M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Teato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto K., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RA "The complete genome sequence of the gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RT Nature 390:249-256 (1997).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=1168;
 RA MEDLINE=90368558; PubMed=1697575;
 RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.,
 RA "Secretory S complex of *Bacillus subtilis*: sequence analysis and
 RT identity to pyruvate dehydrogenase,"
 RL J. Bacteriol. 172:5052-5063 (1990).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=1168;
 RA MEDLINE=97144523; PubMed=8990290;
 RA Henriques A.O., Bryan E.M., Beall B.W., Moran C.P. Jr.,
 RT "cseI5, cse60, and csk22 are new members of mother-cell-specific
 RT sporulation regulons in *Bacillus subtilis*,"
 RL J. Bacteriol. 179:389-398 (1997).
 DR EMBL; Z99111; CAB13300.1; -;
 DR EMBL; AF012285; AAC24901.1; -;
 DR InterPro; IPR002025; NAD-binding.
 DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF00899; Thif; 1.
 KW Complete proteome.
 SQ SEQUENCE 339 AA; 37517 MW; 3CEFS8307DC5C343 CRC64;

Query Match 94.1%; Score 32; DB 16; Length 339;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
 |||||
 Db 275 SLKRLPK 281

RESULT 6
 Q9CJ31 PRELIMINARY; PRT; 478 AA.
 AC Q9CJ31;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Beta-glucosidase A (EC 3.2.1.21).
 GN BGLS OR L10175.
 OS *Lactococcus lactis* (subsp. *lactis*) (Streptococcus *lactis*).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 NCBI_TaxID=1360;

RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21233186; PubMed=11337471;
 RA Boletín A., Winkler P., Mauger S., Jaillon O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.,
 RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
 RT *lactis* ssp. *lactis* IL1403,"
 RL Genome Res. 11:731-753 (2001).
 DR EMBL; AE006255; AAK04273.1; -;
 DR HSP; P11546; IPEG.
 DR InterPro; IPR001360; Glyco_hydro_1.
 DR Pfam; PF00232; Glyco_hydro_1; 1.
 DR PRINTS; PR00131; GLHYDRLASE1.
 DR ProDom; PD006650; Glyco_hydro_1; 1.
 DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 478 AA; 54720 MW; BE3EC9C088B0BE1 CRC64;
 Query Match 94.1%; Score 32; DB 16; Length 478;
 Best Local Similarity 85.7%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
 |||||
 Db 455 SLKRLPK 461

RESULT 7
 Q8YR1 PRELIMINARY; PRT; 905 AA.
 AC Q8YR1;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative LRR.
 GN OSJNBA0026J14.23.
 OC *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.,
 RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
 RL clone:OSJNBA0026J14.23,"
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF004231; BAB99236.1; -;
 DR Gramene; Q8YR1; -;
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 SQ SEQUENCE 905 AA; 103230 MW; F205F3B63C7484C0 CRC64;

Query Match 94.1%; Score 32; DB 10; Length 905;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
 |||||
 Db 574 SLKRLPK 580

RESULT 8
 Q8YH7 PRELIMINARY; PRT; 1381 AA.
 ID Q8YH7;
 AC Q8YH7;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Regulatory protein.
GN ALL1636.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.,
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003586; BAB78002.1; -
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC.1
DR PRINTS; PR00364; DISEASERISIT.
KW Complete proteome.
SQ SEQUENCE 1381 AA; 152325 MW; 7EB862AFB5C33F93 CRC64;
Query Match 94.1%; Score 32; DB 16; Length 1381;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLKRLPK 7
Db 264 SLKRLPK 270
RESULT 9
Q91R55 PRELIMINARY; PRT; 184 AA.
AC Q91R55;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Matrix 2.
GN M2.
OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=38525;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MN/10;
RA Shin H.J., Nagaraja K.V., Halvorson D.A., Njenga M.K.;
RT "Molecular epidemiology of subgroup C avian pneumoviruses isolated
RT from the United States and comparison with subgroup A and B viruses.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028554; AAK38450.1; -
DR InterPro; IPR000571; Znf CCCH.
DR Pfam; PF00642; Zf-CCCH.1.
DR SMART; SM00356; Znf C3H1.1.
DR SEQUENCE 184 AA; 20864 MW; B2C0103E05011DEF CRC64;
Query Match 91.2%; Score 31; DB 12; Length 184;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLKRLPK 7
Db 146 SLKRLPK 152
RESULT 10
Q9YPO9 PRELIMINARY; PRT; 217 AA.
AC Q9YPO9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OX Cucumovirus.
OX NCBI_TaxID=12313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=M1;
RA Xu Z., Higgins C.M., Chen K., Dietzgen R.G., Zhang Z., Yang L.,
RA Fang X.;
RT "Characterization of peanut stunt virus strains from China indicates a
RT third distinct taxonomic subgroup.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222804; CAA10998.1; -
DR InterPro; IPR000247; Cucumovirus_coat.
DR Pfam; PF00760; Cucumo_coat; 1.
DR PRINTS; PR00222; CUCUMOCOAT.
DR ProDom; PD001284; Cucumovirus_coat; 1.
DR SEQUENCE 217 AA; 23688 MW; 573409BC46AFDB4 CRC64;
Query Match 91.2%; Score 31; DB 12; Length 217;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLKRLPK 7
Db 123 SLKRLPK 129
RESULT 11
Q9YPRO PRELIMINARY; PRT; 217 AA.
AC Q9YPRO;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Coat protein.
OS Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RA Xu Z., Higgins C.M., Chen K., Dietzgen R.G., Zhang Z., Yang L.,
RA Fang X.;
RT "Characterization of peanut stunt virus strains from China indicates a
RT third distinct taxonomic subgroup.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222803; CAA10997.1; -
DR InterPro; IPR000247; Cucumovirus_coat.
DR Pfam; PF00760; Cucumo_coat; 1.
DR PRINTS; PR00222; CUCUMOCOAT.
DR ProDom; PD001284; Cucumovirus_coat; 1.
DR SEQUENCE 217 AA; 23716 MW; 5734110B1C92BF3F CRC64;
Query Match 91.2%; Score 31; DB 12; Length 217;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLKRLPK 7
Db 123 SLKRLPK 129
RESULT 12
Q8DSNO PRELIMINARY; PRT; 263 AA.
ID Q8DSNO;
AC Q8DSNO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Putative enoyl-CoA hydratase.
GN SMU.1746C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RA "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen"; Acad. Sci. U.S.A. 99:14434-14439(2002).
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE015003; AAN59379.1; -.
KW Complete proteome.
SQ SEQUENCE 263 AA; 28927 MW; 9A6E426A99F01C6D CRC64;

Query Match 91.2%; Score 31; DB 16; Length 263;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 95 ALKRLPK 101

RESULT 13
Q9CF10 PRELIMINARY; PRT; 348 AA.
AC Q9CF10
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE P0CD.
GN P0CD.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006060; AAK02344.1; -.
DR HSSP; P23861; IPOY.
DR InterPro; IPR006059; SBP_bac_1.
DR InterPro; IPR001188; Sperm/putr-bndng.
DR Pfam; PF01547; SBP_bac_1; 1.
DR PRINTS; PR00909; SPERMNDNDNG.
KW Complete proteome.
SQ SEQUENCE 348 AA; 39116 MW; 69336BAPDC9A0C5 CRC64;

Query Match 91.2%; Score 31; DB 16; Length 348;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 300 ALKRLPK 306

RESULT 14
Q96XL2 PRELIMINARY; PRT; 509 AA.
ID Q96XL2
AC Q96XL2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)

01-DEC-2001 (TReMBLrel. 19, Last sequence update)
01-DEC-2001 (TReMBLrel. 19, Last annotation update)
Hypothetical protein ST2504.
Sulfolobus tokodaii.
Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
Sulfolobus.
NCBI_TaxID=111955;
[1]
SEQUENCE FROM N.A.
STRAIN=JCM 10545 / 7;
PubMed=1572479;
Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
Oshima T., Kikuchi H.;
"Complete genome sequence of an aerobic thermoacidophilic
Crenarchaeon, Sulfolobus tokodaii strain 7.";
DNA Res. 8:123-140(2001).
EMBL; AP000990; BAB67615.1; -.
Hypothetical protein; Complete proteome.
KW SEQUENCE 509 AA; 59579 MW; E7E285524DBE34D5 CRC64;

Query Match 91.2%; Score 31; DB 17; Length 509;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 72 TLKRLPK 78

RESULT 15
O66983 PRELIMINARY; PRT; 669 AA.
ID O66983
AC O66983;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE ATP-dependent DNA helicase REP.
GN REP OR AQ_793.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., R.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
DR EMBL; AE000708; AAC06949.1; -.
DR HSSP; P56255; 1PJR.
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 1.
DR KW Helicase; Complete proteome.
SQ SEQUENCE 669 AA; 76382 MW; 12CBBFC494F51AA3 CRC64;

Query Match 91.2%; Score 31; DB 16; Length 669;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 417 SLKRLPK 423

Search completed: February 11, 2004, 17:09:30
```

us-09-901-187c-7.rsp

Thu Feb 12 09:07:50 2004

Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-7

Perfect score: 34

Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	91.2	186	1 VMA2 TRTV	P33494 turkey rhin
2	31	91.2	308	1 YC55 CYACA	O19888 cyatidum c
3	30	88.2	142	1 YNEK BACSU	P45711 bacillus su
4	30	88.2	145	1 RS15 XIPMA	P70066 xiphophorus
5	30	88.2	453	1 TRME WIGBR	Q8d319 wigglewort
6	30	88.2	1332	1 XKDO BACSU	P54334 bacillus su
7	29	85.3	279	1 ATPG MYCPN	Q50330 mycoplasma
8	29	85.3	294	1 PS2 MYCPN	P75560 mycoplasma
9	29	85.3	313	1 PDE HAZIN	P44800 haemophilus
10	29	85.3	379	1 CYR6 MOUSE	P18406 mus musculus
11	29	85.3	379	1 CYR6 RAT	Q9es72 rattus norv
12	29	85.3	381	1 CYR6 HUMAN	O00622 homo sapien
13	29	85.3	382	1 FATB UMBCA	Q41635 umbellulari
14	29	85.3	395	1 KIME RAT	P17256 rattus norv
15	29	85.3	515	1 CXAA HUMAN	P57773 homo sapien
16	29	85.3	1536	1 SIN3 YEAST	P22579 saccharomyc
17	29	85.3	2164	1 CCNA MOUSE	P97445 mus musculus
18	29	85.3	2212	1 CCNA RAT	P54282 rattus norv
19	29	85.3	2222	1 CCNA RAT	Q07652 rattus norv
20	29	85.3	2223	1 CCNA DISOM	P56699 discopysge o
21	29	85.3	2259	1 CCAB RABIT	Q02343 cryctolagus
22	29	85.3	2272	1 CCAB MOUSE	Q61290 mus musculus
23	29	85.3	2312	1 CCAB HUMAN	Q15878 homo sapien
24	29	85.3	2326	1 CCAB DISOM	P56698 discopysge o
25	29	85.3	2327	1 CCAB MOUSE	Q55017 mus musculus
26	29	85.3	2336	1 CCAB RAT	Q02294 rattus norv
27	29	85.3	2339	1 CCAB HUMAN	Q00975 homo sapien
28	29	85.3	2339	1 CCAB RABIT	Q05152 cryctolagus
29	29	85.3	2424	1 CCNA RABIT	P27884 cryctolagus
30	29	85.3	2505	1 CCNA HUMAN	Q00555 homo sapien
31	28	82.4	102	1 YW02 XANAC	Q9pnp7 xanthomonas
32	28	82.4	146	1 GSPG VIBCH	P45773 vibrio chol
33	28	82.4	150	1 RL15_RICPR	Q9zcs4 rickettsia

34 28 82.4 174 1 VNSC PHODV P35940 phocine dis
35 28 82.4 214 1 FGF8 CHICK Q90722 gallus gall
36 28 82.4 359 1 ADD2 STRCO Q9x7c2 streptomyc
37 28 82.4 428 1 NH44 CAELB Q22555 caenorhabdi
38 28 82.4 435 1 RPOD_SORBI Q01923 sorghum bic
39 28 82.4 507 1 CBS YEAST P32582 saccharomyc
40 28 82.4 509 1 CHLB MESVI Q9mur9 mesostigma
41 28 82.4 576 1 ACH2 DROME P17644 drosophila
42 28 82.4 686 1 RECQ STAAU Q50581 staphylococ
43 28 82.4 732 1 PRIA ECOLI P17888 escherichia
44 28 82.4 830 1 YBMA SCHPO Q10332 schizosacch
45 28 82.4 1479 1 RPOD_WHEAT Q9xps9 triticum ae

ALIGNMENTS

```

RESULT 1
VMA2 TRTV
ID VMA2 TRTV STANDARD; PRT; 186 AA.
AC P33494;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Matrix glycoprotein M2 (Envelope-associated 22 kDa protein).
GN 22K.
OS Turkey rhinotracheitis virus (TRTV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=11264;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333255; PubMed=1629697;
RA Ling R., Easton A.J., Pringle C.R.;
RT "Sequence analysis of the 22K SH and G genes of turkey
RT rhinotracheitis virus and their intergenic regions reveals a gene
RT order different from that of other pneumoviruses.";
RL J. Gen. Virol. 73:1709-1715(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UK/3BV/85;
RC MEDLINE=92300329; PubMed=1607858;
RA Yu Q., Davis P.J., Brown T.D.K., Cavanagh D.;
RT "Sequence and in vitro expression of the M2 gene of turkey
RT rhinotracheitis pneumovirus.";
RL J. Gen. Virol. 73:1355-1363(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S40185; AAB22544.1; -
CC EMBL; X63408; CAA45004.1; -
CC PIR; JQ1623; JQ1623.
CC PIR; JQ1987; JQ1987.
CC InterPro; IPR000571; Znf CCCH.
CC Pfam; PF00642; ZF-CCCH; 1.
CC SMART; SM00356; Znf C3H1; 1.
CC Matrix protein; Envelope protein; Glycoprotein.
CC CARBOHYD 86 86 N-LINKED (GLCNAC... ) (POTENTIAL).
CC CONFLICT 175 175 N -> S (IN REF. 2).
CC SEQUENCE 186 AA; 20986 MW; BC63B574E612E5 CRC64;

```

Query Match 91.2%; Score 31; DB 1; Length 186;
Best Local Similarity 85.7%; Pred. No. 7.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7

|||||


```

Db          146 SLKRLPR 152

RESULT 2
YC55_CVACA
ID YC55_CVACA STANDARD; PRT; 308 AA.
AC O19888;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 36.5 kDa protein ycf55.
GN YCF55 OR YCF5.
OS Cyanidium caldarium.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KK-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentín K.-U.;
RT "The structure and gene repertoire of an ancient red algal plastid
RL genome.";
RL J. Mol. Evol. 51:382-390 (2000).
CC - SIMILARITY: BELONGS TO THE YCF55 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AF022186; AAB82701.1; -
DR PIR; T11956; T11956.
KW Hypothetical protein; Chloroplast.
SQ SEQUENCE 308 AA; 36498 MW; 758C719E056D859F CRC64;

Query Match          91.2%; Score 31; DB 1; Length 308;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLKRLPK 7
    |||||
Db 258 SLKLPLK 264

RESULT 3
YNEK_BACSU
ID YNEK_BACSU STANDARD; PRT; 142 AA.
AC P45711;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yneK.
GN YNEK.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124194; PubMed=8969507;
RA Rose M., Entian K.D.;
RT "New genes in the 170 degrees region of the Bacillus subtilis genome
RT encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino
RT acid transporter.";
RL Microbiology 142:3097-3101 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;

```

```

RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Comerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lepidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Riolva C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpietra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256 (1997).
RN [3]
RP SEQUENCE OF 41-142 FROM N.A.
RC STRAIN=168;
RX MEDLINE=97221596; PubMed=9068642;
RA Schloett T., von Wachenfeldt C., Hederstedt L.;
RT "Identification and characterization of the cda gene, required for
RT cytochrome c synthesis in Bacillus subtilis.";
RL J. Bacteriol. 179:1962-1973 (1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; Z73234; CAA97597.1; -
DR EMBL; Z59113; CABI3680.1; -
DR EMBL; X87845; CAA61119.1; -
DR PIR; D69891; D69891.
DR Subtilisin; BG11252; yneK.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 142 AA; 16989 MW; 30F697A4A4990295 CRC64;

Query Match          88.2%; Score 30; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LKELPK 7
    |||||
Db 33 LKELPK 38

RESULT 4
RS15_XIPMA
ID RS15_XIPMA STANDARD; PRT; 145 AA.
AC P70066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

DE 40S ribosomal protein S15 (RIG protein).

GN RPS15 OR RIG.

OS Xiphophorus maculatus (Southern platyfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorphia; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Cyprinodontiformes; Poeciliidae; Xiphophorus.

OX NCBI_TaxID=8083;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=JP 163A;

RC Walter R.B., Obermoeller R.D., Moore D.D., Lacson J.M.,

RA Coletta L., McEntire B.B., Morizot D.C., Nairn R.S.;

RL Submitted (NCV-1996) to the EMBL/GenBank/DBJ databases.

CC -/- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; U65408; AAB18956.1; -

DR HSP; P80381; 10KF

DR InterPro; IPR002222; Ribosomal_S19.

DR InterPro; IPR005713; S15_euk_arch.

DR Pfam; PF00203; Ribosomal_S19; 1.

DR PRINTS; PR00975; RIBOSOMALS19.

DR ProDom; PD001012; Ribosomal_S19; 1.

DR TIGRFAMs; TIGR01025; rpsS arch; 1.

DR PROSITE; PS00323; RIBOSOMAL_S19; FALSE_NEG.

KW Ribosomal protein.

SQ SEQUENCE 145 AA; 16946 MW; 15896870D21D5351 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 145;

Best Local Similarity 100.0%; Pred. No. 9.6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

DB 57 LKRLPK 62

RESULT 5

TRME WIGBR

ID TRME WIGBR STANDARD; PRT; 453 AA.

AC QED319;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE tRNA modification GTPase time.

GN TRME OR THDF OR WIGBR0120.

OS Wigglesworthia glossinidia brevipalpis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Wigglesworthia.

OX NCBI_TaxID=36870;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22297718; PubMed=12219091;

RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,

RA Aksoy S.;

RT "Genome sequence of the endocellular obligate symbiont of tsetse

RT flies, Wigglesworthia glossinidia";

RL Nat. Genet. 32:402-407(2002).

CC -/- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.

CC Involved in the biosynthesis of the hypermodified nucleoside 5-

CC methylaminomethyl-2-thiouridine, which is found in the wobble

CC position of some tRNAs (By similarity).

CC -/- SIMILARITY: Belongs to the era/trmE family of GTP-binding

CC proteins. TrmE subfamily.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AB063521; BAC24158.1; -

DR HAMAP; MF 00379; -; 1.

DR InterPro; IPR005289; GTP-binding_dom.

DR InterPro; IPR006073; GTP1_OBG.

DR InterPro; IPR005225; Small_GTP.

DR InterPro; IPR004520; rhdf.

DR PRINTS; PR00326; GTP1_OBG.

DR TIGRFAMs; TIGR00650; MG442; 1.

DR TIGRFAMs; TIGR00231; small_GTP; 1.

DR TIGRFAMs; TIGR00450; thdf; 1.

KW tRNA processing; GTP-binding; Complete proteome.

FT NP_BIND 223 230 GTP (POTENTIAL).

FT NP_BIND 270 274 GTP (POTENTIAL).

FT NP_BIND 333 336 GTP (POTENTIAL).

SQ SEQUENCE 453 AA; 50774 MW; 58C06F24BED7A7CE CRC64;

Query Match 88.2%; Score 30; DB 1; Length 453;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

DB 37 LKRLPK 42

RESULT 6

XKDO_BACSU

ID XKDO_BACSU STANDARD; PRT; 1332 AA.

AC P54334;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phase-like element PBSX protein xkdo.

GN XKDO

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Krogh S., O'Reilly M., Nolan N., Devine K.M.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,

RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

```

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni K.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambitt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wisot A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- SIMILARITY: STRONG, TO B.SUBTILIS YQBO.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z70177; CAA94037.1; -.
CC DR EMBL: Z99110; CAB13125.1; -.
CC DR PIR: F69732; F69732.
CC DR Subtilist; BG11549; xtdO.
CC DR InterPro; IPR000189; SLT_domain.
CC DR Pfam; PF01464; SLT; 1.
CC KW Complete proteome.
CC SQ SEQUENCE 1332 AA; 145149 MW; A96C9C9F9B31DF01 CRC64;
Query Match 88.2%; Score 30; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LKRLPK 7
DB 89 LKRLPK 94
RESULT 7
ATPG_MYCPN
ID ATPG_MYCPN STANDARD; PRT; 279 AA.
AC Q50330.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase gamma chain (EC 3.6.3.14).
DE ATPG OR MPN599 OR MP243.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT cluster of ribosomal protein genes."
RL Nucleic Acids Res. 24:628-639(1996).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=ATCC 29342 / M129;
CC RX MEDLINE=97105885; PubMed=8948633;
CC RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
CC RA Herrmann R.;
CC RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
CC pneumoniae."
CC RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON.
CC -1- GRADIENT ACROSS THE MEMBRANE. THE GAMMA CHAIN IS BELIEVED TO BE
CC IMPORTANT IN REGULATING ATPASE ACTIVITY AND THE FLOW OF PROTONS
CC
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni K.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambitt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wisot A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- SIMILARITY: STRONG, TO B.SUBTILIS YQBO.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U43738; AAC43658.1; -.
CC DR EMBL: AE000024; AAB95891.1; -.
CC DR PIR: S62848; S62848.
CC DR InterPro; IPR000131; ATPase_gamma.
CC DR Pfam; PF00231; ATP-synt; 1.
CC DR PRINTS; PR00126; ATPASEGAMMA.
CC DR TIGRFAMs; TIGR01146; ATPSVN_Flgamma; 1.
CC DR PROSITE; PS00153; ATPASE_GAMMA; FALSE NEG.
CC ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase;
KW Complete proteome.
KW SEQUENCE 279 AA; 32394 MW; D3093D376DE2C89E CRC64;
Query Match 85.3%; Score 29; DB 1; Length 279;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
DB 59 SLKRLPK 65
RESULT 8
RS2_MYCPN
ID RS2_MYCPN STANDARD; PRT; 294 AA.
AC R75560;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2.
DE RPSB OR MPN208 OR MP623.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000059; AAB96271.1; -.
CC DR PIR: S73949; S73949.
CC DR HAWAP; MF_00291; -.
CC DR InterPro; IPR001865; Ribosomal_S2.

```

```

DR InterPro; IPR005706; S2_bact.org.
DR Pfam; PF00318; Ribosomal_S2_1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMs; TIGR01011; rpsB_bact; 1. FALSE NEG.
DR PROSITE; PS00962; RIBOSOMAL_S2_1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
DR RIBOSOMAL protein; Complete proteome.
SQ SEQUENCE 294 AA; 33425 MW; 60D111400CSF350D CRC64;

Query Match 85.3%; Score 29; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLP 6
    |||||
Db 168 SLKRLP 173

RESULT 9
ID PLDB HAEIN STANDARD; PRT; 313 AA.
AC P44800;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable lysophospholipase L2 (EC 3.1.1.5) (Lecithinase B).
GN PLDB OR H10645.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
ON NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Kerslaghe A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
CC glycerophosphocholine + a fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U32747; AAC22305.1; -.
CC PIR; A64084; A64084.
CC TIGR; H10645; -.
CC InterPro; IPR000073; A/b_hydrolase.
CC InterPro; IPR000379; Ser estsr_site.
CC Pfam; PF00561; abhydrolase; 1.
KW Hydrolase; Lipid synthesis; Inner membrane; Complete proteome.
SQ SEQUENCE 313 AA; 36658 MW; 891E784FF7F77C2E CRC64;

Query Match 85.3%; Score 29; DB 1; Length 313;
Best Local Similarity 71.4%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7

```

```

Db 237 AIKSLPK 243
    ::|||
RESULT 10
CYR6 MOUSE
ID CYR6 MOUSE STANDARD; PRT; 379 AA.
AC P18406;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10) (3CH61).
GN CYR61 OR IGFBP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAH5/C; TISSUE=Fibroblast;
RX MEDLINE=90287146; PubMed=2355916;
RA O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
RT "Expression of cyr61, a growth factor-inducible immediate-early gene.";
RT Mol. Cell. Biol. 10:3569-3577(1990).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=AJ; TISSUE=Embryonic fibroblast;
RX MEDLINE=91288203; PubMed=2062642;
RA Latinkic B.V., O'Brien T.P., Lau L.F.;
RT "Promoter function and structure of the growth factor-inducible immediate early gene cyr61.";
RT Nucleic Acids Res. 19:3261-3267(1991).
RL
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
CC IN LUNG.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC -!- INDUCTION: By growth factors.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M32490; AAA37512.1; -.
CC EMBL; X56790; CAA40109.1; -.
CC PIR; A35669; A35669.
CC MGD; MGI:88613; Cyr61.
CC GO; GO:0001569; P:patterning of blood vessels; IMP.
CC InterPro; IPR006208; Cys_knot.
CC InterPro; IPR006207; Cys_knot_C.
CC InterPro; IPR000867; Insi_gro_fac_pr.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR001007; WFCC.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00219; IGFBP; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; wvc_1; 1.
CC SMART; SM00041; C1; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSP1; 1.

```

DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK 1; 1.
DR PFAM; PF00007; Cys_knot; 1.
DR PROSITE; PS01225; IGF_BINDING; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWC; 1.
DR PROSITE; PS0184; VWC; 2; 1.
DR Growth factor binding; Signal.
FT SIGNAL 1 24
FT CHAIN 25 379
FT DOMAIN 98 164
FT DOMAIN 226 271
FT DOMAIN 284 358
FT DISULFID 284 321
FT DISULFID 301 335
FT DISULFID 312 351
FT DISULFID 315 353
FT DISULFID 320 357
SQ SEQUENCE 379 AA; 41709 MW; 6465014B56A8EB9 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
DB 202 SLKRLP 207
|||||

RESULT 11
CYR6_RAT STANDARD; PRT; 379 AA.
AC Q9B572;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (insulin-like growth factor-binding protein 10).
GN CYR61 OR IGFBP10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=20435857; PubMed=10852911;
RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
RA Nitsch R.M.;
RT "Muscarinic acetylcholine receptors induce the expression of the
RT immediate early growth regulatory gene CYR61."
RL J. Biol. Chem. 275:28928-28936(2000).
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 VWC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF218568; AAG14964.1; -;
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.

DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWC_C.
DR PFAM; PF00007; Cys_knot; 1.
DR PFAM; PF00219; IGFBP; 1.
DR PFAM; PF00090; TSP1; 1.
DR PFAM; PF00093; VWC; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK 1; 1.
DR PROSITE; PS01225; CTCK 2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWC; 1; 1.
DR PROSITE; PS0184; VWC; 2; 1.
KW Growth factor binding; Signal.
FT SIGNAL 1 24
FT CHAIN 25 379
FT DOMAIN 98 164
FT DOMAIN 226 271
FT DOMAIN 284 358
FT DISULFID 284 321
FT DISULFID 301 335
FT DISULFID 312 351
FT DISULFID 315 353
FT DISULFID 320 357
SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
DB 202 SLKRLP 207
|||||

RESULT 12
CYR6_HUMAN STANDARD; PRT; 381 AA.
ID Q00622; O14934; Q98ZL7;
AC Q00622; O14934; Q98ZL7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (insulin-like growth factor-binding protein 10) (IGI protein).
GN CYR61 OR IGFBP10 OR GIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97280750; PubMed=9135077;
RA Jay P., Berge-Leffrand J.L., Marsollier C., Mejean C., Taviaux S.,
RA Berta P.;
RT "The human growth factor-inducible immediate early gene, CYR61, maps
RT to chromosome 1p.";
RL Oncogene 14:1753-1757(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98197344; PubMed=9536281;
RA Martinerie C., Vigas-Pequignot E., Nguyen V.C., Perbal B.;
RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
RT cells from the nervous system.";
RL Mol. Pathol. 50:310-316(1997).
RN [4]

RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kolesnikova T.V., Lau L.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Bi A.B., Xu L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Schuetze N., Lechner A., Groll C., Koehrl J., Jakob F.;
RL "Regulation of hCYP61 by vitamin D, serum and cytokines in fetal human
osteoblasts";
RT fibroblasts";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Leng B., Tai G., Eatable M., Liu J., Chow C., Sadowski I.;
RL "Organization and expression of the CYP61 gene in normal human
fibroblasts";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung, Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RL "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 VWC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y12084; CAA72802.1; -;
DR EMBL; U62015; AAB58319.1; -;
DR EMBL; Y11307; CRA72167.1; -;
DR EMBL; AF003594; AAB61240.1; -;
DR EMBL; AF031385; AAB84227.1; -;
DR EMBL; Z98053; CAB10848.1; -;
DR EMBL; AF307860; AAG59863.1; -;
DR EMBL; BC001271; AAH01271.1; -;
DR EMBL; BC009199; AAH09199.1; -;
DR EMBL; BC016952; AAH16952.1; -;
DR EMBL; HGNC:2654; CYP61.
DR Genew; 602369; -;
DR MW; 602369; -;

DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Ins1_gro_fac_pr.
DR InterPro; IPR000884; TSP1_gro_fac_pr.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF02119; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00292; TSP1; 1.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS01184; VWC_2; 1.
KW Growth factor binding; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 381 CYR61 PROTEIN.
FT DOMAIN 26 97 IGFBP.
FT DOMAIN 98 164 VWC.
FT DOMAIN 228 273 TSP TYPE-1.
FT DOMAIN 286 360 CTCK.
FT DISULFID 286 323 BY SIMILARITY.
FT DISULFID 303 337 BY SIMILARITY.
FT DISULFID 314 353 BY SIMILARITY.
FT DISULFID 317 355 BY SIMILARITY.
FT DISULFID 322 359 BY SIMILARITY.
FT CONFLICT 165 215 E -> Q (IN REF. 3).
FT CONFLICT 210 210 L -> I (IN REF. 5).
FT CONFLICT 220 220 L -> R (IN REF. 5).
FT CONFLICT 369 369 F -> L (IN REF. 7).
SQ SEQUENCE 381 AA; 42026 MW; FC0BD39C078CA0B1 CRC64;
Query Match 85.3%; Score 29; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLP 6
DB 206 SLKRLP 211
RESULT 13
FATE UMBCA
ID FATE UMBCA STANDARD; PRT; 382 AA.
AC Q41635;
DT 30-MAY-2000 (Rel. 39; Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Lauroyl-acyl carrier protein thioesterase, chloroplast precursor
DE (EC 3.1.2.14) (12:0-acyl-carrier protein thioesterase) (12:0-ACP
DE thioesterase) (Acyl-[acyl-carrier protein] hydrolase) (BTE).
DE FATS1 OR FATS.
GN Umbellularia californica (California bay).
OS Umbellularia californica; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; Laurales; Umbellales; Umbellulalia.
OC NCBI_TaxID=3438;
OX [1]_UMBCA
RP SEQUENCE FROM N.A.
RX MEDLINE=92320297; PubMed=1621095;
RA Voelker T.A., Worrell A.C., Anderson L., Bleibaum J., Fan C.,
RA Hawkins D.J., Radke S.E., Davies H.M.;
RT "Fatty acid biosynthesis redirected to medium chains in transgenic
RL oilseed plants";
RL Science 257:72-74 (1992).
RN [2]
RN MUTAGENESIS TO CHANGE SPECIFICITY.

```

RX MEDLINE=96068671; PubMed=7479856;
RA Yuan L., Voelker T.A., Hawkins D.J.;
RT "Modification of the substrate specificity of an acyl-acyl carrier
RL protein thioesterase by protein engineering.";1995.
RL Proc. Natl. Acad. Sci. U.S.A. 92:10639-10643(1995).
CC -|- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CHAIN TERMINATION DURING DE
CC NOVO FATTY ACID SYNTHESIS. HIGH THIOESTERASE ACTIVITY FOR LAUROYL-
CC ACP VS. OTHER ACRYL-ACPS.
CC -|- CATALYTIC ACTIVITY: Lauroyl-[acyl-carrier protein] + H(2)O =
CC [acyl-carrier protein] + laurate.
CC -|- SUBCELLULAR LOCATION: Chloroplast.
CC -|- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in
CC canola by Monsanto (Calgene) so as to obtain laurate-rich seeds.
CC -|- SIMILARITY: BELONGS TO THE ACRYL-ACP THIOESTERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M94159; AAA34215.1; -.
CC PIR; A40229; A40229.
CC InterPro; IPR002864; Acyl-ACP TE.
CC Pfam; PF01643; Acyl-ACP TE; 1.
CC KX Fatty acid biosynthesis; Hydrolase; Chloroplast; Transit peptide;
CC Genetically modified food.
CC TRANSIT 1 60
CC CHAIN 61 382
CC MUTAGEN 197 197 M-R: CONVERTED TO A 14:0 ACP TE; WHEN
CC MUTAGEN 199 199 ASSOCIATED WITH H-199 AND K-231.
CC MUTAGEN 231 231 ASSOCIATED WITH R-197 AND K-231.
CC SEQUENCE 382 AA; 42915 MW; 96262A558545E89F CRC64;
CC CHOROPLAST (POTENTIAL).
CC LAUROYL-ACYL CARRIER PROTEIN
CC THIOESTERASE.
CC M-R: CONVERTED TO A 14:0 ACP TE; WHEN
CC ASSOCIATED WITH H-199 AND K-231.
CC R-SH: CONVERTED TO A 14:0 ACP TE; WHEN
CC ASSOCIATED WITH R-197 AND K-231.
CC T-SK: CONVERTED TO A 14:0 ACP TE; WHEN
CC ASSOCIATED WITH R-197 AND H-199.
CC -----
CC Query Match 85.3%; Score 29; DB 1; Length 382;
CC Best Local Similarity 100.0%; Pred. No. 46;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 SLKRLP 6
CC DB 56 SLKRLP 61
CC
CC RESULT 14
CC KIME RAT STANDARD; PRT; 395 AA.
CC ID KIME RAT STANDARD; PRT; 395 AA.
CC AC P17256;
CC DT 01-AUG-1990 (Rel. 15, Created)
CC DT 01-AUG-1990 (Rel. 15, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Mevalonate kinase (EC 2.7.1.36) (MK).
CC GN MVK.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
CC RC STRAIN=Sprague-Dawley; TISSUE=Liver;
CC RX MEDLINE=9022132; PubMed=2158094;
CC RA Tanaka R.D., Lee L.Y., Schafer B.L., Kratunis V.J., Mohler W.A.,
CC Robinson G.W., Mosley S.T.;
CC RA "Molecular cloning of mevalonate kinase and regulation of its mRNA
CC levels in rat liver";
CC Proc. Natl. Acad. Sci. U.S.A. 87:2872-2876(1990).
CC -|- FUNCTION: MAY BE A REGULATORY SITE IN CHOLESTEROL BIOSYNTHETIC
CC PATHWAY.

```

```

CC -|- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-
CC phosphomevalonate.
CC -|- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE
CC COMPETITIVE INHIBITORS.
CC -|- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.
CC -|- SUBUNIT: Homodimer.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.
CC -|- DISEASE: MUTATION IN THE MEVALONATE KINASE GENE CAUSES MEVALONIC
CC ACIDURIA.
CC -|- SIMILARITY: BELONGS TO THE GHMP KINASE FAMILY. MEVALONATE KINASE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29472; AAA41588.1; -.
CC PIR; A35629; A35629.
CC PDB; IKVK; 1S-JUN-02.
CC InterPro; IPR001174; Galkinase.
CC InterPro; IPR006204; GHMP Kinase.
CC InterPro; IPR006203; GHMPKase ATP.
CC InterPro; IPR006205; Mev gal Kin.
CC InterPro; IPR006206; Mev gal Kinase.
CC Pfam; PF00288; GHMP_kinases; 1.
CC PRINTS; PR00960; LMEPPROTEIN.
CC PRINTS; PR00959; MEVGALKINASE.
CC TIGRFAms; TIGR00549; mevalon kin; 1.
CC PROSITE; PS00627; GHMP_KINASES ATP; 1.
CC KW Transferase; Kinase; Cholesterol biosynthesis; ATP-binding;
CC Peroxisome; 3D-structure.
CC NP_BIND 128 148
CC FT NP_BIND 128 148
CC FT SEQUENCE 395 AA; 41987 MW; 803D1F44E3C525FC CRC64;
CC SQ
CC
CC Query Match 85.3%; Score 29; DB 1; Length 395;
CC Best Local Similarity 100.0%; Pred. No. 48;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 SLKRLP 6
CC DB 223 SLKRLP 228
CC
CC RESULT 15
CC CXAA HUMAN STANDARD; PRT; 515 AA.
CC ID CXAA HUMAN STANDARD; PRT; 515 AA.
CC AC P57773;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Gap junction alpha-10 protein (Connexin 59) (Cx59).
CC GN GJA10.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA Coucke P.J., Van Laer L., Meyers J., Van Hauwe P., Ottschytch N.,
CC Wauters G., Kelley P., Willems P.J., Van Camp G.;
CC RT "Identification of a new connexin gene using degenerate PCR
CC primers";
CC RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.
CC -|- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -|- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II)

```

```
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF179597; AAC09406.1; -.
CC Genew; HGNC:19155; GJAL0.
CC InterPro; IPR000500; Connexin.
CC Pfam; PF00029; connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.
CC PROSITE; PS00407; CONNEXINS_1; 1.
CC PROSITE; PS00408; CONNEXINS_2; 1.
CC Gap junction; Transmembrane.
CC DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 20 40 POTENTIAL.
CC DOMAIN 41 77 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 78 98 POTENTIAL.
CC DOMAIN 99 166 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 167 187 POTENTIAL.
CC DOMAIN 188 209 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 210 230 POTENTIAL.
CC DOMAIN 231 515 CYTOPLASMIC (POTENTIAL).
CC SEQUENCE 515 AA; 58856 MW; EEBF0B9510AD0C96 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 515;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
Db 274 SLKRLP 279
```

Search completed: February 11, 2004, 17:04:11
Job time : 6.16667 secs

A:Accession: F69659
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-339 <KUN>
A:Cross-references: GB:Z99111; GB:AL009126; NID:G2633699; PIDN:CAB13300.1; PID:el185017;
A:Experimental source: strain 168
C:Genetics:
A:Gene: moeB

Query Match 94.1%; Score 32; DB 2; Length 339;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||
DB 275 SLKRLPK 281

RESULT 3
G86646
beta-glucosidase (EC 3.2.1.21) [imported] - Lactococcus lactis subsp. lactis (strain IL1
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86646
R:Boletín, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:111337471
A:Accession: G86646
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-478 <STO>
A:Cross-references: GB:AB005176; PID:gl2723027; PIDN:AAK04273.1; GSPDB:GN00146
A:Experimental source: strain il1403
C:Genetics:
A:Gene: bgls
A:Superfamily: Agrobacterium beta-glucosidase
C:Keywords: glycosidase; hydrolase

Query Match 94.1%; Score 32; DB 2; Length 478;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||
DB 455 SLKRLPK 461

RESULT 4
AF2010
regulatory protein all1636 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AF2010
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurihara, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1381 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB78002.1; PID:gl17135456; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1636

Query Match 94.1%; Score 32; DB 2; Length 1381;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||
DB 264 SLKRLPK 270

RESULT 5
JQ1623
envelope-associated 22K protein - turkey rhinotracheitis virus
C:Species: turkey rhinotracheitis virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: JQ1623
R:Jilling, R.; Easton, A.J.; Pringle, C.R.
J. Gen. Virol. 73, 1709-1715, 1992
A:Title: Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus a
A:Reference number: PQ0405; MUID:92333255; PMID:1629697
A:Accession: JQ1623
A:Molecule type: mRNA
A:Residues: 1-186 <LIN>
A:Cross-references: GB:S40185; NID:G251600; PIDN:AAE22544.1; PID:G251602
C:Genetics:
A:Gene: 22K
C:Superfamily: respiratory syncytial virus envelope-associated 22K protein
C:Keywords: glycoprotein
F:86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 1; Length 186;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||
DB 146 SLKRLPK 152

RESULT 6
JQ1987
hypothetical 21.0K protein - turkey rhinotracheitis virus (strain UK/3BV/85)
N:Alternate names: ORF 1 protein
C:Species: turkey rhinotracheitis virus
C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
C:Accession: JQ1987
R:Yu, Q.; Davis, P.J.; Brown, T.D.K.; Cavanagh, D.
J. Gen. Virol. 73, 1355-1363, 1992
A:Title: Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pneu
A:Reference number: JQ1987; MUID:92300329; PMID:1607858
A:Accession: JQ1987
A:Molecule type: mRNA
A:Residues: 1-186 <YUQ>
A:Cross-references: GB:X63408; NID:G297846; PIDN:CAA45004.1; PID:G297847
C:Superfamily: respiratory syncytial virus envelope-associated 22K protein

Query Match 91.2%; Score 31; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||
DB 146 SLKRLPK 152

RESULT 7
T11956
hypothetical protein ORF308 - red alga (Cyanidium caldarium) chloroplast
C:Species: Chloroplast Cyanidium caldarium
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 05-May-2000
C:Accession: T11956
R:Gloeckner, G.; Rosenthal, A.; Valentin, K.
submitted to the EMBL Data Library, September 1997
A:Description: Organisation of 46 kb of the Cyanidium caldarium RK1 plasmid genome.
A:Reference number: Z17374
A:Accession: T11956
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A;Residues: 1-308 <GLO>
 A;Cross-references: EMBL:AF022186; NID:g2465730; PIDN:AA82701.1; PID:g2465772
 A;Experimental source: strain RK1
 C;Genetics:
 A;Genome: chloroplast
 A;Note: ycf5
 C;Keywords: chloroplast

Query Match 91.2%; Score 31; DB 2; Length 308;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 258 SLKRLPK 264

RESULT 8

D70369 ATP-dependent DNA helicase REP - Aquifex aeolicus

C;Species: Aquifex aeolicus
 C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C;Accession: D70369
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: D70369

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-669 <AOF>

A;Cross-references: GB:AE000708; NID:g2983356; PIDN:AC06949.1; PID:g2983362; GB:AE00065

A;Experimental source: strain VF5

C;Genetics:

A;Gene: rep

C;Superfamily: helicase II

Query Match 91.2%; Score 31; DB 2; Length 669;
 Best Local Similarity 85.7%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 417 SLKRLPK 423

RESULT 9

D69891 yneK protein - Bacillus subtilis

C;Species: Bacillus subtilis
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C;Accession: D69891; S57765

R;Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Enian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maues
 Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A;Authors: Schleif, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrak, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: D69891

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-142 <KUN>

A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13680.1; PID:g2634180
 A;Experimental source: strain 168
 R;Schlott, T.; von Wachenfeldt, C.; Hederstedt, L.
 submitted to the EMBL Data Library, June 1995
 A;Reference number: S57401

A;Accession: S57785

A;Molecule type: DNA

A;Residues: 41-142 <SCH>

A;Cross-references: EMBL:X87845; NID:g870921; PIDN:CAA61119.1; PID:g870928

A;Experimental source: strain 168

C;Genetics:

A;Gene: yneK

C;Superfamily: Bacillus subtilis yneK protein

Query Match 88.2%; Score 30; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LKRLPK 7
 |||||
 Db 33 LKRLPK 38

RESULT 10

D83941

hypothetical protein BH2332 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: D83941

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: D83941

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-151 <STO>

A;Cross-references: GB:AF001515; GB:BA000004; NID:gi0174886; PIDN:BA06051.1; GSPDB:GNO

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH2332

C;Superfamily: Bacillus subtilis yneK protein

Query Match 88.2%; Score 30; DB 2; Length 151;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LKRLPK 7
 |||||
 Db 32 LKRLPK 37

RESULT 11

H90209

GTP-binding protein [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C;Accession: H90209

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovici, Y.; Allard, G.; Awayez, M.J.; Chan

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,

arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: H90209

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-332 <KUR>

A;Cross-references: GB:AE006641; NID:gi13813793; PIDN:AAK40935.1; GSPDB:GN00155

C;Genetics:

A;Gene: SSO0625

C;Superfamily: GTP-binding protein, GTP1/OBG; translation elongation factor Tu homology

```

Query Match      88.2%; Score 30; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 49;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
      |||||
Db      22 LKRLPK 27

RESULT 12
efflux protein homolog lin2916 [imported] - Listeria innocua (strain Clp11262)
A:Gene: AE1796
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1796
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <GLA>
A:Cross-references: GB:AL59202; PIDN:CAC98141.1; PID:G16415457; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2916

Query Match      88.2%; Score 30; DB 2; Length 396;
Best Local Similarity 85.7%; Pred. No. 58;
Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      |||||
Db      181 SLKRLPE 187

RESULT 13
efflux protein homolog lmo2777 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1421
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00990.1; PID:G16412277; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2777

Query Match      88.2%; Score 30; DB 2; Length 396;
Best Local Similarity 85.7%; Pred. No. 58;
Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      |||||
Db      181 SLKRLPE 187

```

```

RESULT 14
A:Gene: AE2817
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82817
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-735 <STM>
A:Cross-references: GB:AE003887; GB:AE003849; NID:g9105175; PIDN:AAF83162.1; GSPDB:GN00
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, P.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0352
C:Superfamily: guanosine 3',5'-bis(diphosphate) 3'-pyrophosphatase

Query Match      88.2%; Score 30; DB 2; Length 735;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      |||||
Db      532 SLERLPK 538

RESULT 15
Probable serine/threonine kinase P23O10.20 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96716
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: B96716
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <STO>
A:Cross-references: GB:AE005173; NID:g7705100; PIDN:AAF67779.1; GSPDB:GN00141
C:Genetics:
A:Gene: P23O10.20
A:Map position: 1

Query Match      88.2%; Score 30; DB 2; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 LKRLPK 7
| | | | |
Db 81 LKRLPK 86

Search completed: February 11, 2004, 17:11:41
Job time : 10.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 801455 seqs, 209382283 residues 801455
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	7	10	US-09-901-187B-7
2	32	94.1	478	12	US-10-369-493-18305
3	31	91.2	45	11	US-09-764-831-3818
4	31	91.2	669	12	US-10-369-493-584
5	30	88.2	62	10	US-09-764-877-1778
6	30	88.2	62	12	US-10-242-515-1778
7	30	88.2	219	16	US-10-176-306-40
8	29	85.3	178	12	US-10-108-260A-3713
9	29	85.3	271	9	US-09-815-242-5295
10	29	85.3	273	8	US-08-781-986A-5196
11	29	85.3	362	15	US-10-100-121-20
12	29	85.3	374	9	US-09-853-625B-12
13	29	85.3	375	11	US-09-901-910-7
14	29	85.3	379	9	US-09-853-625B-11
15	29	85.3	379	9	US-09-853-625B-11

16	29	85.3	379	12	US-10-099-322-45	Sequence 45, Appl
17	29	85.3	379	12	US-10-182-432-2	Sequence 2, Appl
18	29	85.3	379	12	US-10-044-564-45	Sequence 45, Appl
19	29	85.3	379	14	US-10-053-753-2	Sequence 2, Appl
20	29	85.3	381	11	US-09-901-910-2	Sequence 2, Appl
21	29	85.3	381	12	US-10-394-015-5	Sequence 5, Appl
22	29	85.3	381	12	US-10-099-322-42	Sequence 42, Appl
23	29	85.3	381	12	US-10-099-322-43	Sequence 43, Appl
24	29	85.3	381	12	US-10-099-322-44	Sequence 44, Appl
25	29	85.3	381	12	US-10-182-432-4	Sequence 42, Appl
26	29	85.3	381	12	US-10-044-564-42	Sequence 44, Appl
27	29	85.3	381	12	US-10-044-564-43	Sequence 43, Appl
28	29	85.3	381	14	US-10-044-564-44	Sequence 4, Appl
29	29	85.3	381	14	US-10-053-753-4	Sequence 2, Appl
30	29	85.3	381	15	US-10-284-796-2	Sequence 10, Appl
31	29	85.3	382	15	US-10-100-121-10	Sequence 1432, Ap
32	29	85.3	455	9	US-09-925-301-1432	Sequence 18471, A
33	29	85.3	477	12	US-10-369-493-18471	Sequence 47, Appl
34	29	85.3	515	12	US-09-863-776-47	Sequence 8, Appl
35	29	85.3	522	12	US-10-029-120-8	Sequence 8, Appl
36	29	85.3	522	14	US-10-027-806-8	Sequence 8, Appl
37	29	85.3	522	15	US-10-034-623-8	Sequence 8, Appl
38	29	85.3	522	15	US-10-027-801-8	Sequence 2, Appl
39	29	85.3	538	15	US-10-013-803A-2	Sequence 3953, Ap
40	29	85.3	546	12	US-10-094-749-1953	Sequence 3683, Ap
41	29	85.3	582	12	US-10-108-260A-3683	Sequence 12128, A
42	29	85.3	642	12	US-10-369-493-12128	Sequence 18, Appl
43	29	85.3	1478	12	US-10-334-143-18	Sequence 322, App
44	29	85.3	1536	10	US-09-901-368-322	Sequence 5496, Ap
45	29	85.3	2198	12	US-10-369-493-5496	

ALIGNMENTS

RESULT 1

US-09-901-187B-7
Sequence 7, Application US/09901187B
Patent No. US20020151464A1

GENERAL INFORMATION:
APPLICANT: Panacea Pharmaceuticals, Inc.

APPLICANT: Wolozin, Benjamin

APPLICANT: Ostretova-Golts, Natalie

APPLICANT: Lebowitz, Michael S.

TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment

TITLE OF INVENTION: Alpha-Synuclein Diseases

FILE REFERENCE: PAN01/002US

CURRENT APPLICATION NUMBER: US/09/901,187B

CURRENT FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/217,319

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: US 60/279,199

PRIOR FILING DATE: 2001-03-28

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 7

TYPE: PRT

ORGANISM: Homo sapiens

US-09-901-187B-7

Query Match 100.0%; Score 34; DB 10; Length 7;

Best Local Similarity 100.0%; Pred. No. 7e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7

Db 1 SLKRLPK 7

RESULT 2

US-10-369-493-18305

Sequence 18305, Application US/10369493

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18305
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Lactococcus lactis
; US-10-369-493-18305

Query Match      94.1%; Score 32; DB 12; Length 478;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      455 SLKRIPK 461

RESULT 3
US-09-764-891-3818
; Sequence 3818, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3818
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-3818

Query Match      91.2%; Score 31; DB 11; Length 45;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      29 ALKRLPK 35

RESULT 4
US-10-369-493-58
; Sequence 58, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18305
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Lactococcus lactis
; US-10-369-493-18305

Query Match      91.2%; Score 31; DB 12; Length 669;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      417 SLKKLPK 423

RESULT 5
US-09-764-877-1778
; Sequence 1778, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1778
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-877-1778

Query Match      88.2%; Score 30; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
DB      20 LKRLPK 25

RESULT 6
US-10-242-515-1778
; Sequence 1778, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
```

```

; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1778
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-242-515-1778

Query Match      88.2%; Score 30; DB 12; Length 62;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
Db      20 LKRLPK 25

RESULT 7
US-10-176-306-40
; Sequence 40, Application US/10176306
; Publication No. US20030130485A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20

```

```

; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-176-306-40

Query Match      88.2%; Score 30; DB 16; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
Db      168 LKRLPK 173

RESULT 8
US-10-108-260A-3713
; Sequence 3713, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3713
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3713

Query Match      85.3%; Score 29; DB 12; Length 178;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLP 6
Db      139 SLKRLP 144

RESULT 9
US-09-815-242-5295
; Sequence 5295, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```



```

/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 5295
/ LENGTH: 271
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-5295

Query Match      85.3%; Score 29; DB 9; Length 271;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      122 SLKRLPK 128

RESULT 10
US-08-781-986A-5196
/ Sequence 5196, Application US/08781986A
/ Publication No. US20030054436A1
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5255
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,986A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 5196:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 273 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-781-986A-5196

Query Match      85.3%; Score 29; DB 8; Length 273;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      122 SLKRLPK 128

```

```

RESULT 11
US-09-815-242-12548
/ Sequence 12548, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12548
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12548

Query Match      85.3%; Score 29; DB 9; Length 273;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      122 SLKRLPK 128

RESULT 12
US-10-100-121-20
/ Sequence 20, Application US/10100121
/ Publication No. US20030097696A1
/ GENERAL INFORMATION:
/ APPLICANT: Knauf, Vic C
/ APPLICANT: Thompson, Gregory
/ TITLE OF INVENTION: PLANT SEED OILS
/ FILE REFERENCE: MTC 6823.1
/ CURRENT APPLICATION NUMBER: US/10/100,121
/ CURRENT FILING DATE: 2002-08-16
/ PRIOR APPLICATION NUMBER: US 08/926,522
/ PRIOR FILING DATE: 1997-09-09
/ PRIOR APPLICATION NUMBER: US 08/458,173
/ PRIOR FILING DATE: 1995-06-02
/ PRIOR APPLICATION NUMBER: US 07/949,102
/ PRIOR FILING DATE: 1992-09-21
/ PRIOR APPLICATION NUMBER: US 07/762,762
/ PRIOR FILING DATE: 1991-09-16
/ PRIOR APPLICATION NUMBER: WO/US91/01746
/ PRIOR FILING DATE: 1991-03-14
/ PRIOR APPLICATION NUMBER: PCT/US91/05801
/ PRIOR FILING DATE: 1991-08-15

```

;; PRIOR APPLICATION NUMBER: US 07/615,784
;; PRIOR FILING DATE: 1990-11-14
;; NUMBER OF SEQ ID NOS: 48
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 20
;; LENGTH: 362
;; TYPE: PRT
;; ORGANISM: Brassica campestris
US-10-100-121-20

Query Match 85.3%; Score 29; DB 15; Length 362;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 235 SLKXPK 241

RESULT 13

US-09-853-625B-12
; Sequence 12, Application US/09853625B
; Patent No. US20020049304A1

;; GENERAL INFORMATION:
;; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
;; TITLE OF INVENTION: Human CCN-Like Growth Factor
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 INCH DISKETTE
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: WORD PERFECT 5.1
;; CURRENT APPLICATION NUMBER: US/09/853,625B
;; FILING DATE: 14-May-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION NUMBER: 09/053,587
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MULLINS, J.G.
;; REGISTRATION NUMBER: 33,073
;; REFERENCE/DOCKET NUMBER: 325800-442
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 374 AMINO ACIDS
;; TYPE: AMINO ACID
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: PROTEIN
;; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-853-625B-12

Query Match 85.3%; Score 29; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
DB 207 SLKRLP 212

RESULT 14

US-09-901-910-7
; Sequence 7, Application US/09901910
; Publication No. US20030012768A1
;; GENERAL INFORMATION:
;; APPLICANT: Li, Hacdong
;; APPLICANT: Adams, Mark
;; APPLICANT: Calenda Valerie
;; TITLE OF INVENTION: Connective Tissue Growth Factor-2
;; FILE REFERENCE: PFI26P2
;; CURRENT APPLICATION NUMBER: US/09/901,910
;; CURRENT FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 09/348,815
;; PRIOR FILING DATE: 1999-07-08
;; PRIOR APPLICATION NUMBER: 08/459,101
;; PRIOR FILING DATE: 1995-06-02
;; PRIOR APPLICATION NUMBER: PCT/US94/07736
;; PRIOR FILING DATE: 1994-07-12
;; PRIOR APPLICATION NUMBER: 60/217,402
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/291,642
;; PRIOR FILING DATE: 2001-05-18
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 7
;; LENGTH: 375
;; TYPE: PRT
;; ORGANISM: homo sapiens
US-09-901-910-7

Query Match 85.3%; Score 29; DB 11; Length 375;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
DB 207 SLKRLP 212

RESULT 15

US-09-853-625B-11
; Sequence 11, Application US/09853625B
; Patent No. US20020049304A1
;; GENERAL INFORMATION:
;; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
;; TITLE OF INVENTION: Human CCN-Like Growth Factor
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 INCH DISKETTE
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: WORD PERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/853,625B
;; FILING DATE: 14-May-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION NUMBER: 09/053,587
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MULLINS, J.G.
;; REGISTRATION NUMBER: 33,073
;; REFERENCE/DOCKET NUMBER: 325800-442
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700

```
;
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 379 AMINO ACIDS
;   TYPE: AMINO ACID
;   STRANDEDNESS: <Unknown>
;   TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-853-625B-11

Query Match      85.3%  Score 29; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLXRLP 6
Db      202 SLXRLP 207

Search completed: February 11, 2004, 17:54:10
Job time : 25.6667 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-7

Perfect score: 34

Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	3118	4	US-09-579-181-1
2	30	88.2	145	4	US-09-732-210-1428
3	29	85.3	139	4	US-09-252-991A-24681
4	29	85.3	319	2	US-07-745-206A-17
5	29	85.3	319	2	US-08-311-363-17
6	29	85.3	362	1	US-08-464-523B-32
7	29	85.3	362	2	US-08-948-176-25
8	29	85.3	362	2	US-08-440-845D-6
9	29	85.3	362	3	US-08-868-458-6
10	29	85.3	362	4	US-09-303-592-2
11	29	85.3	362	4	US-09-303-592-4
12	29	85.3	366	2	US-08-605-106-11
13	29	85.3	367	1	US-08-075-533-2
14	29	85.3	367	2	US-08-948-176-2
15	29	85.3	367	5	PCT-US91-09160-2
16	29	85.3	370	1	US-07-662-007B-37
17	29	85.3	370	1	US-07-824-247-37
18	29	85.3	370	3	US-08-470-204A-37
19	29	85.3	374	1	US-08-468-847B-12
20	29	85.3	375	2	US-08-459-101A-2
21	29	85.3	379	1	US-08-468-847B-11
22	29	85.3	379	4	US-09-142-569-2
23	29	85.3	381	4	US-09-142-569-4
24	29	85.3	381	4	US-09-348-845-2
25	29	85.3	382	1	US-07-824-247-42
26	29	85.3	382	1	US-08-142-473A-2
27	29	85.3	382	1	US-08-464-523B-26

28	29	85.3	382	1	US-08-469-203A-2	Sequence 2, Appli
29	29	85.3	382	1	US-08-469-203A-2	Sequence 2, Appli
30	29	85.3	382	2	US-08-605-106-12	Sequence 12, Appli
31	29	85.3	382	2	US-08-440-845D-1	Sequence 1, Appli
32	29	85.3	382	3	US-08-470-204A-42	Sequence 42, Appli
33	29	85.3	382	3	US-08-868-458-1	Sequence 1, Appli
34	29	85.3	591	4	US-09-252-991A-26716	Sequence 26716, A
35	29	85.3	1182	3	US-09-041-886-21	Sequence 21, Appli
36	29	85.3	1754	1	US-07-745-206A-13	Sequence 13, Appli
37	29	85.3	1754	2	US-08-311-363-13	Sequence 13, Appli
38	29	85.3	2237	1	US-08-455-543A-48	Sequence 48, Appli
39	29	85.3	2237	2	US-08-223-305C-48	Sequence 8, Appli
40	29	85.3	2237	4	US-09-268-163-8	Sequence 36, Appli
41	29	85.3	2265	2	US-08-149-097D-36	Sequence 36, Appli
42	29	85.3	2336	4	US-09-268-163-10	Sequence 10, Appli
43	29	85.3	2337	3	US-08-713-118-2	Sequence 2, Appli
44	29	85.3	2337	3	US-09-452-007-2	Sequence 2, Appli
45	29	85.3	2339	1	US-08-455-543A-47	Sequence 47, Appli

ALIGNMENTS

RESULT 1
US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chrivia, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1

Query Match 100.0%; Score 34; DB 4; Length 3118;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 69 SLKRLPK 75

RESULT 2
US-09-732-210-1428
; Sequence 1428, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1428

```

; LENGTH: 145
; TYPE: PRT
; ORGANISM: Xiphophorus maculatus
; US-09-732-210-1428

Query Match      88.2%; Score 30; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
DB      57 LKRLPK 62

RESULT 3
US-09-252-991A-24681
; Sequence 24681, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24681
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24681

Query Match      85.3%; Score 29; DB 4; Length 139;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      67 SLRLPK 73

RESULT 4
US-07-745-206A-17
; Sequence 17, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815

Query Match      85.3%; Score 29; DB 2; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-745-206A-17

Query Match      85.3%; Score 29; DB 1; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
DB      23 TIKRLPK 29

RESULT 5
US-08-311-363-17
; Sequence 17, Application US/08311363
; Patent No. 5876358
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-311-363-17

Query Match      85.3%; Score 29; DB 2; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SLKRLPK 7
Db 23 TLKRLPK 29

RESULT 6
US-08-464-523B-32
; Sequence 32, Application US/08464523B
; Patent No. 5723761
; GENERAL INFORMATION:
; APPLICANT: Toni A. Voelker
; APPLICANT: Ling Yuan
; APPLICANT: Jean Kridl
; APPLICANT: Deborah Hawkins
; APPLICANT: Aubrey Jones
; TITLE OF INVENTION: Plant Acyl ACP Thioesterase
; TITLE OF INVENTION: Sequences
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,523B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13131
; FILING DATE: 10-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,004
; FILING DATE: 10-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 100-1WO
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-523B-32

Query Match 85.3%; Score 29; DB 1; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 235 SLKKIPK 241

RESULT 8
US-08-440-845D-6
; Sequence 6, Application US/08440845D
; Patent No. 5955329
; GENERAL INFORMATION:
; APPLICANT: Yuan, L.
; APPLICANT: Kridl, J.
; APPLICANT: Knauf, V.
; APPLICANT: Dehesh, K.
; TITLE OF INVENTION: Engineering Plant Thioesterases For
; TITLE OF INVENTION: Altered Substrate Specificity.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
; COMPUTER: Apple Macintosh

QY 1 SLKRLPK 7
Db 23 TLKRLPK 29

RESULT 6
US-08-464-523B-32
; Sequence 32, Application US/08464523B
; Patent No. 5723761
; GENERAL INFORMATION:
; APPLICANT: Toni A. Voelker
; APPLICANT: Ling Yuan
; APPLICANT: Jean Kridl
; APPLICANT: Deborah Hawkins
; APPLICANT: Aubrey Jones
; TITLE OF INVENTION: Plant Acyl ACP Thioesterase
; TITLE OF INVENTION: Sequences
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,523B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13131
; FILING DATE: 10-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,004
; FILING DATE: 10-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 100-1WO
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-523B-32

Query Match 85.3%; Score 29; DB 1; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 235 SLKKIPK 241

RESULT 7
US-08-948-176-25
; Sequence 25, Application US/08948176

```

; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,845D
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-845D-6

Query Match      85.3%; Score 29; DB 2; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 9
US-08-868-458-6
; Sequence 6, Application US/08868458
; Patent No. 6150512
; GENERAL INFORMATION:
; APPLICANT: Yuan, L.
; TITLE OF INVENTION: Engineering Plant Thioesterases And
; Disclosure of Plant Thioesterases
; Having No. 6150512el Substrate Specificity
;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: Microsoft Word For Window 95 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,458
; FILING DATE: 03-Jun-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07064
; FILING DATE: 15-MAY-96
; APPLICATION NUMBER: 08/537,083
; FILING DATE: 29-SEPT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-845D-6

Query Match      85.3%; Score 29; DB 3; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 10
US-09-303-592-2
; Sequence 2, Application US/09303592
; Patent No. 6426448
; GENERAL INFORMATION:
; APPLICANT: BOOTH, JOHN R.
; APPLICANT: BROGLIE, RICHARD M.
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: KINNEY, ANTHONY J.
; APPLICANT: KNOWLTON, SUSAN
; APPLICANT: SEBASTIAN, SCOTT A.
; APPLICANT: E.I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: NOVEL GENE COMBINATIONS THAT ALTER THE QUALITY AND
; FILE REFERENCE: BB-1156
; CURRENT APPLICATION NUMBER: US/09/303,592
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 60/085,030
; EARLIER FILING DATE: MAY 11, 1998
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 2
; LENGTH: 362
; TYPE: PPT
; ORGANISM: Brassica napus
; US-09-303-592-2

Query Match      85.3%; Score 29; DB 4; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 11
US-09-303-592-4
; Sequence 4, Application US/09303592
; Patent No. 6426448
; GENERAL INFORMATION:
; APPLICANT: BOOTH, JOHN R.
; APPLICANT: BROGLIE, RICHARD M.
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: KINNEY, ANTHONY J.
; APPLICANT: KNOWLTON, SUSAN
; APPLICANT: SEBASTIAN, SCOTT A.
; APPLICANT: E.I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: NOVEL GENE COMBINATIONS THAT ALTER THE QUALITY AND
; FILE REFERENCE: BB-1156
; CURRENT APPLICATION NUMBER: US/09/303,592
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 60/085,030
; EARLIER FILING DATE: MAY 11, 1998
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 4
```

```
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-303-592-4

Query Match      85.3%; Score 29; DB 4; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 12
US-08-605-106-11
; Sequence 11, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235.001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-106-11

Query Match      85.3%; Score 29; DB 2; Length 366;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLKRLPK 7
Db      234 SLKKIPK 240

RESULT 13
US-08-075-533-2
; Sequence 2, Application US/08075533
; Patent No. 5530166
```

```
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D.
; APPLICANT: Yadav, Narendra S.
; TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
; TITLE OF INVENTION: Thioesterase Genes
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,533
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631,264
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Morrissey, Bruce W.
; REGISTRATION NUMBER: 30,663
; REFERENCE/DOCKET NUMBER: CR-8926-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4927
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-075-533-2

Query Match      85.3%; Score 29; DB 1; Length 367;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLKRLPK 7
Db      232 SLKKIPK 238

RESULT 14
US-08-948-176-2
; Sequence 2, Application US/08948176
; Patent No. 5945585
; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: YADAV, NARENDRA S.
; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES
; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT
; TITLE OF INVENTION: OIL COMPOSITION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0
```


Thu Feb 12 09:07:48 2004

PCT-US91-09160-2

Query Match 85.3%; Score 29; DB 2; Length 367;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
DB 232 SLKKIPK 238

Search completed: February 11, 2004, 17:13:37
Job time : 11.4167 secs

QY 1 SLKRLPK 7
DB 232 SLKKIPK 238

RESULT 15

PCT-US91-09160-2

Sequence 2, Application PC/TUS9109160

GENERAL INFORMATION:

APPLICANT: Hitz, William D.

APPLICANT: Yadav, Narendra S.

TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP

TITLE OF INVENTION: Thioesterase Genes

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. du Pont de Nemours and Company

STREET: 1007 Market Street

CITY: Wilmington

STATE: Delaware

COUNTRY: U.S.A

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/09160

FILING DATE: 19911216

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/631,264

FILING DATE: 20-DEC-1990

ATTORNEY/AGENT INFORMATION:

NAME: Morrissey, Bruce W.

REGISTRATION NUMBER: 30,663

REFERENCE/DOCKET NUMBER: CR-8926-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 992-4927

TELEFAX: (302) 892-7949

TELEX: 835420

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-7

Sequence: 1 SLKRLPX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	7	23 AAE14552	Human alpha-synuclein
2	34	100.0	3118	22 AAE50362	Human SRCAP. Homo
3	32	94.1	478	23 ABE33479	Lactococcus lactis
4	31	91.2	45	22 ABB95861	Human cesticular a
5	31	91.2	45	22 AAM95160	Human reproductive
6	30	88.2	62	22 ABB03831	Human musculoskele
7	30	88.2	62	22 ABU13125	Novel human muscu
8	30	88.2	219	23 AAU99919	Human 49875 DEAD t
9	30	88.2	254	24 ABP78272	N. gonorrhoeae ami

10	30	88.2	396	23 ABB49928	Listeria monocytog
11	30	88.2	448	22 ARG24660	Novel human diagno
12	30	88.2	966	22 ABB68113	Drosophila melanog
13	29	85.3	102	23 ABP09012	Human ORFX protein
14	29	85.3	165	23 ABG06433	Arabidopsis thalia
15	29	85.3	165	23 ABG06070	Human DITP Polype
16	29	85.3	171	21 AAG54995	Arabidopsis thalia
17	29	85.3	174	21 AAG14734	Arabidopsis thalia
18	29	85.3	217	21 AAG06432	Arabidopsis thalia
19	29	85.3	229	21 AAG06431	Arabidopsis thalia
20	29	85.3	229	21 AAG39849	Arabidopsis thalia
21	29	85.3	230	21 AAG14733	Arabidopsis thalia
22	29	85.3	235	21 AAG39886	Arabidopsis thalia
23	29	85.3	271	22 AAU33799	Staphylococcus aur
24	29	85.3	273	18 AAW89748	Staphylococcus aur
25	29	85.3	273	22 AAU36955	Arabidopsis thalia
26	29	85.3	291	21 AAG39985	Arabidopsis thalia
27	29	85.3	306	21 AAG14732	Arabidopsis thalia
28	29	85.3	362	13 AAR29172	Brassica thioester
29	29	85.3	362	15 AAR54948	Brassica campestri
30	29	85.3	362	16 AAR74150	Brassica campestri
31	29	85.3	362	19 AAW44333	Brassica napus acy
32	29	85.3	362	20 AAY28638	Brassica rapa C18:
33	29	85.3	362	20 AAW87467	Arabidopsis thalia
34	29	85.3	362	21 AAG39848	Arabidopsis thalia
35	29	85.3	362	21 AAY53760	An acyl-ACP thioes
36	29	85.3	362	21 AAY53761	An acyl-ACP thioes
37	29	85.3	362	23 ABB97569	Acyl acyl carrier
38	29	85.3	362	23 ABB92431	Herbicide daily activ
39	29	85.3	366	14 AAR41673	Rape acyl-ACP thio
40	29	85.3	366	14 AAR41674	Rape acyl-ACP thio
41	29	85.3	367	13 AAR28339	Acyl-ACP thioester
42	29	85.3	367	20 AAY28832	Soybean seed acyl-
43	29	85.3	367	21 AAG39984	Arabidopsis thalia
44	29	85.3	367	23 ABB92915	Herbicide daily activ
45	29	85.3	370	12 AAR14802	Bay thioesterase.

ALIGNMENTS

RESULT 1

AAE14552
ID AAE14552 standard; peptide; 7 AA.

AC AAE14552;

XX 17-MAY-2002 (first entry)

DT Human alpha-synuclein aggregation inhibitor #7.

DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.

OS Homo sapiens.

XX WO200204482-A1

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21379.

XX 07-JUL-2000; 2000US-217319P.

XX 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

XX Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

PT Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper
 PS Claim 40; Page 37; 52pp; English.
 XX
 CC The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 34; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 1 SLKRLPK 7

RESULT 2
 AAB50362
 ID AAB50362 standard; protein; 3118 AA.

AC AAB50362;
 DT 12-MAR-2001 (first entry)
 DE Human SRCAP.

Human; SRCAP; Snf2 related CBP activator protein; antiviral; CREB;
 CAMP regulatory element; CREB binding protein; CBP; ATPase;
 transcription activation; DEAD box RNA dependent helicase;
 KW adenoviral DBP protein; beta-actin; nuclear receptor; viral infection.
 XX Homo sapiens.

WO200073467-A1.
 07-DEC-2000.

25-MAY-2000; 2000WO-US14719.
 27-MAY-1999; 99US-0136620.
 25-MAY-2000; 2000US-0579181.

(UYSL-) UNIV SAINT LOUIS.
 Chrivia J, Yaciuk P;

WPI; 2001-061545/07.
 DR N-PSDB; AAC89859.

Snf2 related CAMP regulatory element (CREB) binding protein (CBP)
 activator protein, capable of co-activating CREB binding protein,
 useful for modulating transcription and for affecting viral infection -

Claim 6; Page 77-86; 103pp; English.

The present sequence is an Snf2 related CREB (CAMP regulatory element)
 binding protein (CBP) activator protein (SRCAP) polypeptide. It has
 ATPase activity and is capable of activating transcription. SRCAP
 polypeptides are useful for activating transcription in a cell, for
 enhancing CREB (CAMP regulatory element) binding protein (CBP)-mediated
 activation of transcription in a cell, for treating a patient having a
 disease involving a function such as insufficient transcription of a

CC gene, e.g. a gene mediated by CBP co-activation, DEAD box RNA dependent
 CC helicase, adenoviral DBP protein, beta-actin or a nuclear receptor
 CC affected by SRCAP protein. Compounds that modulate SRCAP function, such
 CC as antibodies, antisense molecules, polynucleotides or ribozymes, are
 CC useful for treating diseases mediated by SRCAP-activated transcription,
 CC for example, infection by adenovirus, hepatitis C virus, human
 CC immunodeficiency virus type-1, Epstein Barr virus, cytomegalovirus or
 CC hepatitis B virus.
 XX

SQ Sequence 3118 AA;

Query Match 100.0%; Score 34; DB 22; Length 3118;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||
 Db 69 SLKRLPK 75

RESULT 3
 ABB53479
 ID ABB53479 standard; protein; 478 AA.

AC ABB53479;
 DT 16-MAY-2002 (first entry)
 DE Lactococcus lactis protein bgls.

Lactococcus lactis protein bgls.
 Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX Lactococcus lactis IL1403.

FR2807446-A1.
 12-OCT-2001.

11-APR-2000; 2000FR-0004630.
 11-APR-2000; 2000FR-0004630.

(INRG) INRA INST NAT RECH AGRONOMIQUE.
 Bolotine A, Sorokine A, Renault P, Ehrlich SD;

WPI; 2002-043418/06.
 New nucleotide sequence useful in the identification or Lactococcus
 lactis and related species -

Claim 6; SEQ ID No 181; 2504pp; French.

The present invention is related to a Lactococcus lactis nucleotide
 sequence (A90521) and related proteins (ABB53300-ABB55621). The
 nucleic acid sequence is useful in the detection and/or amplification of
 nucleic acid sequence, particularly to identify Lactococcus lactis or
 related species. The proteins of the invention are useful for the
 biosynthesis or biodegradation of a composition of interest. The
 invention helps research in lactic bacteria, particularly useful in the
 production of yogurt and cheese.

Note: The sequence data for this patent is based on equivalent patent
 WO200177334 (published 18-Oct-2001) which is available in electronic
 format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 478 AA;

Query Match 94.1%; Score 32; DB 23; Length 478;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
 |||||

PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254037.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 XX WPI; 2001-483232/52.
 DR
 XX
 XX Nucleic acids encoding 973 human testicular antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating testicular cancer -
 XX
 XX Claim 11; SEQ ID NO 1245; 766pp; English.
 XX
 XX The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a protein of the
 CC invention.
 XX
 XX SQ Sequence 45 AA;
 Query Match 91.2%; Score 31; DB 22; Length 45;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SLKRLPK 7
 :|||||
 Db 29 ALKRLPK 35
 RESULT 5
 ID AAM95160
 ID AAM95160 standard; Protein; 45 AA.
 AC AAM95160;
 XX
 XX DT 21-NOV-2001 (first entry)
 DE
 DE Human reproductive system related antigen SEQ ID NO: 3818.
 XX
 XX Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy.
 KW
 KW Homo sapiens.
 OS
 XX WO200155320-A2.
 PN
 XX
 XX PD 02-AUG-2001.
 XX
 XX PF 17-JAN-2001; 2001WO-US01339.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.

CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 XX
 SQ Sequence 45 AA;
 Query Match 91.2%; Score 31; DB 22; Length 45;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SLKRLPK 7
 DB 29 ALKRLPK 35
 RESULT 6
 ABB03831
 ID ABB03831 standard; Protein; 62 AA.
 XX AC ABB03831;
 XX
 XX 08-JAN-2002 (first entry)
 DE Human musculoskeletal system related polypeptide SEQ ID NO 1778.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system.
 XX
 OS Homo sapiens.
 XX
 PN WO200155367-A1.
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01338.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226682.
 PR 22-AUG-2000; 2000US-0226683.
 PR 22-AUG-2000; 2000US-0226684.
 PR 22-AUG-2000; 2000US-0226685.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0226687.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0226689.
 PR 22-AUG-2000; 2000US-0226690.
 PR 22-AUG-2000; 2000US-0226691.
 PR 22-AUG-2000; 2000US-0226692.
 PR 22-AUG-2000; 2000US-0226693.
 PR 22-AUG-2000; 2000US-0226694.
 PR 22-AUG-2000; 2000US-0226695.
 PR 22-AUG-2000; 2000US-0226696.
 PR 22-AUG-2000; 2000US-0226697.
 PR 22-AUG-2000; 2000US-0226698.
 PR 22-AUG-2000; 2000US-0226699.
 PR 22-AUG-2000; 2000US-0226700.
 PR 22-AUG-2000; 2000US-0226701.
 PR 22-AUG-2000; 2000US-0226702.
 PR 22-AUG-2000; 2000US-0226703.
 PR 22-AUG-2000; 2000US-0226704.
 PR 22-AUG-2000; 2000US-0226705.
 PR 22-AUG-2000; 2000US-0226706.
 PR 22-AUG-2000; 2000US-0226707.
 PR 22-AUG-2000; 2000US-0226708.
 PR 22-AUG-2000; 2000US-0226709.
 PR 22-AUG-2000; 2000US-0226710.
 PR 22-AUG-2000; 2000US-0226711.
 PR 22-AUG-2000; 2000US-0226712.
 PR 22-AUG-2000; 2000US-0226713.
 PR 22-AUG-2000; 2000US-0226714.
 PR 22-AUG-2000; 2000US-0226715.
 PR 22-AUG-2000; 2000US-0226716.
 PR 22-AUG-2000; 2000US-0226717.
 PR 22-AUG-2000; 2000US-0226718.
 PR 22-AUG-2000; 2000US-0226719.
 PR 22-AUG-2000; 2000US-0226720.
 PR 22-AUG-2000; 2000US-0226721.
 PR 22-AUG-2000; 2000US-0226722.
 PR 22-AUG-2000; 2000US-0226723.
 PR 22-AUG-2000; 2000US-0226724.
 PR 22-AUG-2000; 2000US-0226725.
 PR 22-AUG-2000; 2000US-0226726.
 PR 22-AUG-2000; 2000US-0226727.
 PR 22-AUG-2000; 2000US-0226728.
 PR 22-AUG-2000; 2000US-0226729.
 PR 22-AUG-2000; 2000US-0226730.
 PR 22-AUG-2000; 2000US-0226731.
 PR 22-AUG-2000; 2000US-0226732.
 PR 22-AUG-2000; 2000US-0226733.
 PR 22-AUG-2000; 2000US-0226734.
 PR 22-AUG-2000; 2000US-0226735.
 PR 22-AUG-2000; 2000US-0226736.
 PR 22-AUG-2000; 2000US-0226737.
 PR 22-AUG-2000; 2000US-0226738.
 PR 22-AUG-2000; 2000US-0226739.
 PR 22-AUG-2000; 2000US-0226740.
 PR 22-AUG-2000; 2000US-0226741.
 PR 22-AUG-2000; 2000US-0226742.
 PR 22-AUG-2000; 2000US-0226743.
 PR 22-AUG-2000; 2000US-0226744.
 PR 22-AUG-2000; 2000US-0226745.
 PR 22-AUG-2000; 2000US-0226746.
 PR 22-AUG-2000; 2000US-0226747.
 PR 22-AUG-2000; 2000US-0226748.
 PR 22-AUG-2000; 2000US-0226749.
 PR 22-AUG-2000; 2000US-0226750.
 PR 22-AUG-2000; 2000US-0226751.
 PR 22-AUG-2000; 2000US-0226752.
 PR 22-AUG-2000; 2000US-0226753.
 PR 22-AUG-2000; 2000US-0226754.
 PR 22-AUG-2000; 2000US-0226755.
 PR 22-AUG-2000; 2000US-0226756.
 PR 22-AUG-2000; 2000US-0226757.
 PR 22-AUG-2000; 2000US-0226758.
 PR 22-AUG-2000; 2000US-0226759.
 PR 22-AUG-2000; 2000US-0226760.
 PR 22-AUG-2000; 2000US-0226761.
 PR 22-AUG-2000; 2000US-0226762.
 PR 22-AUG-2000; 2000US-0226763.
 PR 22-AUG-2000; 2000US-0226764.
 PR 22-AUG-2000; 2000US-0226765.
 PR 22-AUG-2000; 2000US

PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

N-PSDB; AAL35413.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Claim 11; SEQ ID NO 1778; 781pp + Sequence Listing; English.

XX The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (antagonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 62 AA;

Query Match 88.2%; Score 30; DB 22; Length 62;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LKRLPK 7

Db 20 LKRLPK 25

RESULT 7
 ABU13125
 ID ABU13125 standard, Protein; 62 AA.
 XX AC
 XX ABU13125;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE Novel human musculoskeletal system antigen #745.
 XX
 XX Musculoskeletal system antigen; cancer; metastasis;
 KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KW post-operative tissue repair; limb regeneration; neuronal growth;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW AIDS-related complex; chondrocyte growth; bone regeneration;
 KW periodontal regeneration; tissue transport; bone graft; skin aging;
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KW cell growth; organ transplant; cell differentiation; body height;
 KW weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KW pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;
 KW depression; tendency for violence; pain; reproductive capability;
 KW hormone level; endocrine level; appetite; libido; memory; stress;
 KW storage capability; fat content; lipid content; protein content;
 KW carbohydrate content; vitamin content; cofactor content;
 KW nutritional component.
 XX
 XX Homo sapiens.
 OS
 XX
 PN US2002147140-A1.
 XX
 PD 10-OCT-2002.
 XX
 XX 17-JAN-2001; 2001US-0764877.
 XX
 XX 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 05-SEP-2000; 2000US-229503P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.

PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 13-OCT-2000; 2000US-237040P.
 PR 20-OCT-2000; 2000US-239935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249239P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX (ROSE/) ROSEN C A.
 FA (RUBE/) RUBEN S M.
 FA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 XX WPI; 2003-128199/12.
 DR N-PSDB; AX56401.
 XX
 XX Isolated nucleic acid molecules encoding musculoskeletal system
 PT associated polypeptides, useful for detecting disorders, e.g. cancer -
 PT Claim 11; SEQ ID NO 1778; 321pp; English.
 PS
 XX The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid: stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, haematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's metal state or physical state by influencing biorhythms.
 CC caradiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This is the amino acid sequence of a
 CC novel human musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.sepdata.uspto.gov/sequence.html?DocID=20020147140.

XX Sequence 62 AA;

Query Match 88.2%; Score 30; DB 24; Length 62;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

|||||

Db 20 LKRLPK 25

AAU99919 standard; Protein; 219 AA.

AAU99919:

07-OCT-2002 (first entry)

Human 49875 DEAD type helicase domain consensus sequence (SMART).

49875; DEAD type helicase domain; infection;
 haematopoietic disorder; blood clotting disorder; cancer;
 autoimmune disorder; leukaemia; immunological disorder;
 cardiovascular disorder; neurological disorder; cellular proliferation;
 red blood cell disorder; viral disease; neurological disorder.

Synthetic.

WO200240656-A2.

23-MAY-2002.

14-NOV-2001; 2001WO-US45291.

14-NOV-2000; 2000US-248331P.

14-NOV-2000; 2000US-248362P.

14-NOV-2000; 2000US-248365P.

30-NOV-2000; 2000US-250077P.

30-NOV-2000; 2000US-250176P.

30-NOV-2000; 2000US-250327P.

(MILL-) MILLENNIUM PHARM INC.

Meyers RE, Curtis RAJ, Glucksmann MA;
 WPI; 2002-508325/54.

Isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041
 or 84234 polypeptides, useful as reagents or targets for treating or
 diagnosing pain or metabolic, liver, kidney, or cardiovascular
 disorders

Disclosure; Fig 8; 298pp; English.

This invention relates to the DNA and protein sequences of novel
 isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041
 or 84234 proteins. The method of the invention is useful for treating a
 disorder characterised by aberrant activity of 47476, 67210, 49875, 46842,
 33201, 83378, 84233, 64708, 85041 or 84234-expressing cell in a
 subject. The protein molecules can act as novel diagnostic targets and
 therapeutic agents for controlling aberrant or deficient signal
 transduction resulting, in e.g., haematopoietic disorders, including
 blood clotting disorders, autoimmune disorders, or disorders related to
 an inability to clear infections (e.g., viral or bacterial infections),
 as well as disorders related to abnormal cellular proliferation or
 differentiation, e.g., leukaemia. They may also be used to control
 disorders of metabolic imbalance (e.g., disorders of lipopolysaccharide
 biosynthesis or glycogen synthesis) immunological disorders,
 cardiovascular disorders, neurological disorders, or cellular
 proliferation and/or differentiation disorders, e.g., cancer, cell
 motility and adhesion disorders, differentiative disorders, red blood
 cell disorders, viral diseases, neurological disorders (e.g., brain
 disorders), pain or metabolic disorders, liver disorders, kidney
 disorders, disorders of the small intestine, disorders of metal ion
 imbalance, protein trafficking disorders and disorders associated with
 bone metabolism. The sequences of the invention are also useful for
 screening assays, predictive medicine (e.g., diagnostic assays,
 prognostic assays, monitoring clinical trials, and pharmacogenetics),
 and methods of treatment (e.g., therapeutic and prophylactic). The
 present sequence represents a predicted consensus sequence motif

CC found in the protein of the invention.

XX Sequence 219 AA;

Query Match 88.2%; Score 30; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7
 Db 168 LKRLPK 173

RESULT 9

ABP78272

ID ABP78272 standard; Protein; 254 AA.

XX AC ABP78272;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 3074.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB02069.

XX PR 12-FEB-2001; 2001GB-0003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizsa M, Masignani V, Monaci E;
 WPI; 2003-058415/05.
 N-PSDB; ABZ39242.

DR New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 medicament for treating or preventing N. gonorrhoeae infection

PT Disclosure; Page 412; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
 Also disclosed are the nucleic acid molecules encoding the proteins and
 antibodies that specifically bind to the proteins. The composition
 comprising the protein, nucleic acid or antibody is useful for the
 manufacture of a medicament for treating or preventing N. gonorrhoeae
 infection, this may be in the form of a vaccine or gene therapy.
 Sequences given in records ABP76736-ABP81046 represent nucleic acid
 molecules of the invention.

XX Sequence 254 AA;

Query Match 88.2%; Score 30; DB 24; Length 254;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
 Db 45 SLKRLPE 51

RESULT 10

ABB49928

ID ABB49928 standard; Protein; 396 AA.

XX AC ABB49928;

XX

DT 05-FEB-2002 (first entry)
 XX Listeria monocytogenes protein #2632.
 DE
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 XX vitamin B12; bacterial infection; disease.
 KW
 XX Listeria monocytogenes.
 OS
 XX
 XX WO200177335-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX
 XX 11-APR-2001; 2001WO-FK01118.
 PF
 XX 11-APR-2000; 2000FR-0004629.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Rueniok C, Feihl H, Dehoux P;
 PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Carrido-Garcia P, Tierrez-Martinez A, Amend L;
 PI Chakraborty T, Donann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 PI
 XX WPI; 2002-010914/01.
 DR
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 XX and prevention of Listeria and related bacterial infections, and
 XX related polypeptides -
 PT
 XX
 PS Claim 6; SEQ ID NO 2633; 192pp; French.
 XX
 XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 396 AA;
 SQ
 Query Match 88.2%; Score 30; DB 23; Length 396;
 Best Local Similarity 85.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SLKRLPK 7
 DB 181 SLKRLPE 187
 RESULT 11
 ABB24660
 ID ABB24660 standard; Protein; 448 AA.
 XX
 AC ABB24660;
 XX
 DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #24651.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-FSDB; AAS8864.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT
 XX Claim 20; SEQ ID NO 55019; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (II) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABB00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 448 AA;
 SQ
 Query Match 88.2%; Score 30; DB 22; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LKRLPK 7
 DB 134 LKRLPK 139
 RESULT 12
 ABB68113
 ID ABB68113 standard; Protein; 966 AA.
 XX
 AC ABB68113;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 31131.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX N-PSDB; ABL12216.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Disclosure; SEQ ID NO 31131; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (AB5737-AB572072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 966 AA;
XX Query Match 88.2%; Score 30; DB 22; Length 966;
XX Best Local Similarity 85.7%; Pred. No. 6.5e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 SLKRLPK 7
XX 228 SLKQLPK 234
XX
XX RESULT 13
XX ID ABP09012 standard; Protein; 102 AA.
XX AC ABP09012;
XX 24-JUN-2002 (first entry)
XX Human ORFX protein sequence SEQ ID NO:18006.
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX Homo sapiens.
XX OS
XX WO200192523-A2.
XX
XX 06-DEC-2001.
XX 29-MAY-2001; 2001WO-US10836.
XX 30-MAY-2000; 2000US-206132P.
XX 29-AUG-2000; 2000US-228716P.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach MD;
XX WPI; 2002-106308/14.
XX N-PSDB; AEN24764.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders -
XX Disclosure; SEQ ID 18006; 1037pp; English.
XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). ABL15762 to ABL27252 encode the human ORFX
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX disorder in humans, and in the manufacture of a medicament for treating a
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX osteoarthritis, neurodegenerative diseases, diabetes mellitus, systemic
XX transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX storage disease, various immune deficiencies and disorders, infectious
XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX bone degenerative disorders, or periodontal disease, and for gut
XX protection or regeneration and treatment of lung or liver fibrosis.
XX reperfusion injury in various tissues and conditions resulting from
XX systemic cytokine damage.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Query Match 85.3%; Score 29; DB 23; Length 102;
XX Best Local Similarity 71.4%; Pred. No. 1.4e+02;
XX Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 SLKRLPK 7
XX 65 SLKQLPK 71
XX
XX RESULT 14
XX ID AAG06433 standard; Protein; 165 AA.
XX AC AAG06433;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3204.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX OS

Thu Feb 12 09:07:48 2004

us-09-901-187c-7.rag

XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0141842.
PN		PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PP		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
PP		PR	13-JUL-1999;	99US-0143542.
XX		PR	14-JUL-1999;	99US-0143624.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005.
PR	05-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	23-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	PR	22-JUL-1999;	99US-0145132.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	19-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	20-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	10-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	11-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	PR	07-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.

```
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 08-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.3%; Score 29; DB 21; Length 165;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 39 SLKRLPK 45

RESULT 15
ABG60070
ID ABG60070 standard; Protein; 165 AA.
XX AC ABG60070;
XX 30-JUL-2002 (first entry)
XX Human DITHP polypeptide #128.
XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;
XX Cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
XX inflammatory disorder; viral infection; bacterial infection; seizure;
XX fungal infection; parasitic infections; developmental disorder; breast;
XX endocrine disorder; metabolic disorder; neurological disorder; cervix;
XX gastrointestinal disorder; transport disorder; gene therapy; kidney;
XX adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
XX thymus.
OS Homo sapiens.
XX
PN WO200220754-A2.
```

```
XX 14-MAR-2002.
XX 29-AUG-2001; 2001WO-US27127.
XX 05-SEP-2000; 2000US-229747P.
XX 05-SEP-2000; 2000US-229748P.
XX 05-SEP-2000; 2000US-229749P.
XX 05-SEP-2000; 2000US-229750P.
XX 05-SEP-2000; 2000US-229751P.
XX 05-SEP-2000; 2000US-230583P.
XX 06-SEP-2000; 2000US-230585P.
XX 06-SEP-2000; 2000US-230514P.
XX 06-SEP-2000; 2000US-230515P.
XX 06-SEP-2000; 2000US-230517P.
XX 06-SEP-2000; 2000US-230518P.
XX 06-SEP-2000; 2000US-230519P.
XX 06-SEP-2000; 2000US-230595P.
XX 06-SEP-2000; 2000US-230597P.
XX 06-SEP-2000; 2000US-230598P.
XX 06-SEP-2000; 2000US-230599P.
XX 06-SEP-2000; 2000US-230610P.
XX 06-SEP-2000; 2000US-230865P.
XX 06-SEP-2000; 2000US-230988P.
XX 07-SEP-2000; 2000US-230951P.
XX 07-SEP-2000; 2000US-231163P.
XX 07-SEP-2000; 2000US-231167P.
XX (INCY-) INCYTE GENOMICS INC.
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
XX Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
XX Momiyama MG, Bradley DU, Rohatgi SD, Harris B, Roseberry AM;
XX Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
XX Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI: 2002-383054/41.
XX N-PSDB; ABK71661.
XX An isolated polynucleotide useful in diagnostics and therapeutics -
XX Claim 29; Page 606; 686pp; English.
XX The invention relates to human diagnostic and therapeutic (dithp)
XX polynucleotides and their associated polypeptides (DITHP polypeptides).
XX The sequences of the invention are used in the treatment and diagnosis of
XX cell proliferative disorders (e.g. atherosclerosis, cirrhosis) cancers
XX (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast, or
XX cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis, or
XX thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
XX psoriasis, osteoporosis), viral infections, bacterial infections, fungal
XX infections, parasitic infections, developmental disorders (e.g. anaemia,
XX epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
XX endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
XX (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
XX amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
XX disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
XX (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
XX ABG59943-ABG60220 represent human DITHP polypeptides of the invention.
XX
XX Qy 1 SLKRLPK 6
XX Db 130 SLKRLPK 135

Query Match 85.3%; Score 29; DB 23; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 6
Db 130 SLKRLPK 135

Search completed: February 11, 2004, 17:02:51
Job time : 33.25 secs
```

us-09-901-187c-7.rag

Thu Feb 12 09:07:48 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL 23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_plant:*
 - 10: sp_phage:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	483	5 Q95SP1	Q95sp1 drosophila
2	36	100.0	533	5 Q9VHV3	Q9vnh3 drosophila
3	33	91.7	812	11 Q35483	Q35483 mus musculus
4	32	88.9	438	2 Q8K2P3	Q8kzp3 thermus the
5	31	86.1	75	16 Q8V8P8	Q8v8f8 listeria mo
6	31	86.1	145	16 Q8VP84	Q8vp84 anabaena sp
7	31	86.1	153	16 Q8DKN7	Q8dkn7 synecococc
8	31	86.1	195	10 Q9C9A5	Q9c9a5 arabidopsis
9	31	86.1	213	10 Q91556	Q91556 arabidopsis
10	31	86.1	280	16 Q8V4N0	Q8v4n0 listeria in
11	31	86.1	281	16 Q928M7	Q928m7 listeria in
12	31	86.1	358	2 Q8GP85	Q8gp85 streptococc
13	31	86.1	384	5 Q9V629	Q9v629 drosophila
14	31	86.1	386	16 Q8VWS7	Q8vws7 anabaena sp
15	31	86.1	391	2 Q9RMT7	Q9rmt7 streptomyc
16	31	86.1	415	16 Q9FC33	Q9fc33 streptomyc

17	31	86.1	491	5 Q9U622	Q9u622 drosophila
18	31	86.1	493	2 Q8GMH4	Q8gmh4 streptomyc
19	31	86.1	1747	10 Q8LHA4	Q8lha4 oryza sativ
20	30	83.3	84	2 Q8KN68	Q8kn68 pseudomonas
21	30	83.3	273	16 Q8PDE6	Q8pde6 xanthomonas
22	30	83.3	329	16 Q8P4X4	Q8p4x4 xanthomonas
23	30	83.3	404	16 Q923E5	Q923e5 xanthomonas
24	30	83.3	404	16 Q87641	Q87641 xanthomonas
25	30	83.3	404	16 Q8PL87	Q8pl87 xanthomonas
26	30	83.3	404	16 Q8FL76	Q8fl76 xanthomonas
27	30	83.3	412	16 Q8PRM5	Q8prm5 xanthomonas
28	30	83.3	412	5 Q962C3	Q962c3 caenorhabdi
29	30	83.3	440	5 Q19730	Q19730 caenorhabdi
30	30	83.3	449	8 Q9TMO9	Q9tm09 cyanidium c
31	30	83.3	607	10 Q9FL24	Q9fl24 arabidopsis
32	30	83.3	1254	5 Q9U308	Q9u308 caenorhabdi
33	29	80.6	165	10 Q899U1	Q899u1 rhizobium 1
34	29	80.6	170	16 Q9XHP7	Q9xhp7 ceratopteris
35	29	80.6	170	16 Q8YN03	Q8yn03 anabaena sp
36	29	80.6	184	2 Q45156	Q45156 bacillus ps
37	29	80.6	203	16 Q9X109	Q9x109 thermotoga
38	29	80.6	210	10 Q9SE48	Q9se48 oryza sativ
39	29	80.6	210	10 Q9SE47	Q9se47 oryza sativ
40	29	80.6	216	11 Q8CE99	Q8cb99 mus musculu
41	29	80.6	231	10 Q8LHJ7	Q8lhj7 oryza sativ
42	29	80.6	245	16 Q8CKZ5	Q8ckz5 versinia pe
43	29	80.6	251	11 Q8ZFG2	Q8zfg2 versinia pe
44	29	80.6	255	11 Q9D2K3	Q9d2k3 mus musculu
45	29	80.6	274	16 Q8V1F3	Q8v1f3 raistonia s

ALIGNMENTS

RESULT 1
Q95SP1 PRELIMINARY; PRT; 483 AA.
ID Q95SP1
AC Q95SP1
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE GH11496P.
GN CG7800.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Parfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060678; AAL28226.1;
DR FlyBase; FBGN003752; CG7800.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 6.
DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
DR PROSITE; PS05056; LRR_TYPICAL; 1.
SQ SEQUENCE 483 AA; 54650 MW; 459439E8FE17F44 CRC64;

Query Match 100.0%; Score 36; DB 5; Length 483;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
|||||

Qy 1 RLGRNQ 7
| | | | |
Db 503 RLGRNQ 509

RESULT 3
ID Q35483 PRELIMINARY; PRT; 812 AA.
AC Q35483;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kruppel-related zinc finger protein.
GN ZFP316 OR EMZF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=yolk sac;
RA Rothstein J.L., Bratford P.;
RT "Murine Emz1 gene: Expression During Development.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031955; AAB87452.1; -
DR HSSP; P08045; 1ZNF.
DR MGD; MGI:1860402; Zfp316.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; Zf-C2H2; 14.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 14.
KW Metal-binding; Zinc; Zinc-finger
SQ SEQUENCE 812 AA; 88817 MW; 8F89DE55FD063DCB CRC64;

Query Match 91.7%; Score 33; DB 11; Length 812;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7
| | | | |
Db 722 RVRGRNQ 728

RESULT 4
ID Q8KZP3 PRELIMINARY; PRT; 438 AA.
AC Q8KZP3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preprotein translocase SecY subunit.
GN SECY.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8;
RA Masui R., Inoue Y., Shibata T., Miki K., Yokoyama S., Kuramitsu S.;
RT "preprotein translocase SecY subunit";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086887; BAC01134.1; -
DR InterPro; IPR002208; SecY.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00344; secy; 1.
DR PRINTS; PR00303; SECYTRNLCASE.
DR TIGRfams; TIGR00967; 3a0501s007; 1.
DR PROSITE; PS00756; SECY_2; 1.

Db 453 RLGRNQ 459

RESULT 2
ID Q9VHV3 PRELIMINARY; PRT; 533 AA.
AC Q9VHV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG7800 protein.
GN CG7800.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fester C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.J., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum X.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AB036678; AAF54196.2; -
DR FlyBase; FBgn0037552; CG7800.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 5.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SW00369; LRR_TYP; 1.
DR PROSITE; PS0506; LRR_TYPICAL; 1.
SQ SEQUENCE 533 AA; 60296 MW; 3532E77037F67A6E CRC64;

Query Match 100.0%; Score 36; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

DR PROSITE; PS00430; TONB DEPENDENT REC 1; 1;
SQ SEQUENCE 438 AA; 48204 MW; 71BBEF926123A692 CRC64;

Query Match
88.9%; Score 32; DB 2; Length 438;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 432 RLGRNR 438

RESULT 5
Q8Y8P8 PRELIMINARY; PRT; 75 AA.
AC Q8Y8P8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein lmo0946.
GN LMO0946.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BAQUERO P., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC99024.1; -
DR ListList; LMO0946; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8810 MW; 8242CB314143B336 CRC64;

Query Match
86.1%; Score 31; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6
Db 70 RLGRNR 75

RESULT 6
Q8Y8P8 PRELIMINARY; PRT; 145 AA.
AC Q8Y8P8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein All4315.
GN All4315.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003596; BAB76014.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 145 AA; 15732 MW; D16371D9C4EBA237 CRC64;

Query Match
86.1%; Score 31; DB 16; Length 145;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6
Db 99 RLGRNR 104

RESULT 7
Q8DKN7 PRELIMINARY; PRT; 153 AA.
AC Q8DKN7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tlx0822 protein.
GN Tlx0822.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Kato H., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005371; BAC08373.1; -
KW Complete proteome.
SQ SEQUENCE 153 AA; 16164 MW; 01B4E4268EDBFC90 CRC64;

Query Match
86.1%; Score 31; DB 16; Length 153;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6
Db 44 RLGRNR 49

RESULT 8
Q9C9A5 PRELIMINARY; PRT; 195 AA.
AC Q9C9A5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative transcription factor (Atlg71030/F23N20_2).
GN F23N20.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=1130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

```

White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Etgu P., Felblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzalli A.,
 Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 Pail G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 Nature 408:816-820(2000).
 [2]
 SEQUENCE FROM N.A.
 Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
 Ban J., Bowser L., Carninci P., Dale J.M., Gibson H.A., Jones T.,
 Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
 Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
 Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 Davis R.W., Theologis A., Ecker J.R.;
 "Arabidopsis cDNA clones.";
 Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [3]
 SEQUENCE FROM N.A.
 Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 Palm C.J., Bowser L., Jones T., Ban J., Chen H., Cheuk R.,
 Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
 Yamada K., Ecker J., Theologis A., Davis R.W.;
 Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 [4]
 SEQUENCE FROM N.A.
 Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Ban J.,
 Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 Ecker J.R.;
 "Arabidopsis ORF clones.";
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 [5]
 SEQUENCE FROM N.A.
 Haas B.J., Volfovsky N., Town C.D., Trukhan M., Alexandrov N.,
 Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 "Full-length messenger RNA sequences greatly improve genome
 annotation.";
 Genome Biol. 0:0-0(2002).
 [6]
 SEQUENCE FROM N.A.
 Brover V., Trukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 Feldmann K.;
 "Full-length cDNA from Arabidopsis thaliana.";
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC016972; AAG51688.1; -;
 DR EMBL; AY048231; AAK82494.1; -;
 DR EMBL; AF370609; AAK43928.1; -;
 DR EMBL; AY01697; AAM0296.1; -;
 DR EMBL; AY087093; AAM64653.1; -;
 DR HSP; P06876; IMEX.
 DR InterPro; IPR001005; Myb DNA binding.
 DR Pfam; PF00249; myb DNA-binding; 1.
 DR PROSITE; PS00334; MYB_3; 1.
 DR PROSITE; PS00090; MYB_3; 1.
 SEQUENCE 195 AA; 22098 MW; B270E341A015967 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 195;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RLGRN 6
 |||||
 DB 22 RLGRN 27
 RESULT 9
 Q39156 PRELIMINARY; PRT; 213 AA.
 ID Q39156
 AC Q39156
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Myb-related transcription factor.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Siliques;
 RX MEDLINE=97149286; PubMed=8996094;
 RA Kirik V., Baumelein H.;
 RT "A novel leaf-specific myb-related protein with a single binding
 repeat";
 RL Gene 183:109-113(1996).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
 DR EMBL; Z68157; CAA92280.1; -;
 DR HSP; P06876; IMEX.
 DR InterPro; IPR001005; Myb DNA binding.
 DR Pfam; PF00249; myb DNA-binding; 1.
 DR SMART; SM00717; SANT; 1.
 DR PROSITE; PS00334; MYB_2; 1.
 DR PROSITE; PS00090; MYB_3; 1.
 DR DNA-binding; Nuclear protein.
 SK QY SEQUENCE 213 AA; 24693 MW; 9DBB050F5FE40ADC CRC64;
 Query Match 86.1%; Score 31; DB 10; Length 213;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RLGRN 6
 |||||
 DB 22 RLGRN 27
 RESULT 10
 Q8Y4N0 PRELIMINARY; PRT; 280 AA.
 ID Q8Y4N0
 AC Q8Y4N0
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein lmo2406.
 GN LMO2406.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OC NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Saquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Paravar A., Dehoux P.,
 Domant E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

```

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Macdueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001)
DR EMBL; AL591983; CAB00484.1; -
DR ListLust; LMO02406; -
DR InterPro; IPR002763; DUF72.
DR Pfam; PF01904; DUF72; 1.
DR ProDom; PD117970; DUF72; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 280 AA; 32186 MW; 7C154678230D125C CRC64;

Query Match 86.1%; Score 31; DB 16; Length 280;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 192 RLGRNQ 198

RESULT 11
Q928M7 PRELIMINARY; PRT; 281 AA.
AC Q928M7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein lin2505.
LN LIN2505.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
SEQUENCE FROM N.A.
STRAIN=CLIP 11262 / Serovar 6a;
RX PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Anand A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshni H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596172; CAC97732.1; -
DR ListLust; LIN02505; -
DR InterPro; IPR002763; DUF72.
DR Pfam; PF01904; DUF72; 1.
DR ProDom; PD117970; DUF72; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 32372 MW; 5FCBACFF0ECC6D87 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 281;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 193 RLGRNQ 199

RESULT 12
Q9GP85

```

```

ID Q9GP85 PRELIMINARY; PRT; 358 AA.
AC Q9GP85
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Eps9L.
LN Eps9L.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
SEQUENCE FROM N.A.
RA Rallu F., Ehrlich D.S., Renault P.;
RT "Diversity of eps operons in Streptococcus thermophilus.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF454499; AAN63755.1; -
SQ SEQUENCE 358 AA; 40644 MW; 964CA552A8B16D99 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 358;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 60 RLGRNQ 66

RESULT 13
Q9V629 PRELIMINARY; PRT; 384 AA.
AC Q9V629
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG13190 protein.
LN CG13190.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL; A003824; AAF58610.1; -
 DR FvBase; FBGN0033662; CG13190.
 SQ SEQUENCE 384 AA; 44833 MW; 0FF9AC48BCB1AEC CRC64;

Query Match 86.1%; Score 31; DB 5; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;
 QY 2 LRGRNQ 7
 Db 347 LRGRNQ 352

RESULT 14

Q8YWS7 PRELIMINARY; PRT; 386 AA.
 AC Q8YWS7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein A11523.
 GN A11523.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2159285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yanada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium *Anabaena* sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003586; BAB7889.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 386 AA; 41303 MW; 23B333C3B9B5AF7 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;
 QY 1 LRGRGN 6
 Db 299 LRGRGN 304

RESULT 15

Q9RMT7 PRELIMINARY; PRT; 391 AA.
 AC Q9RMT7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Aminase.
 GN PAB8.
 OS Streptomyces venezuelae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=54571;

RP SEQUENCE FROM N.A.
 RC STRAIN=ISP5230;
 RA Chang Z., Sun Y., Vining L.C.;
 RT "paba and pabb, the second set of p-aminobenzoic acid synthase genes
 in Streptomyces venezuelae ISP5230.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF189258; AAF01062.1; -
 DR HSP; Q06128; 1QDL.
 DR InterPro; IPR005801; Anth synth_chor.
 DR InterPro; IPR006056; YigF-like.
 DR Pfam; PF00425; chorismate bind; 1.
 DR PRINTS; PRO0095; ANTSYNTHASE1.
 DR PRODOM; PD000779; Anth synth_chor; 1.
 DR PROSITE; PS01094; UPF0076; 1.
 SQ SEQUENCE 391 AA; 43299 MW; 64D3207A87B41065 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 391;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;
 QY 1 LRGRGN 6
 Db 127 LRGRGN 132

Search completed: February 11, 2004, 17:09:32
 Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	86.1	185	1 KGUA_FUSNN	Q8rh19 fusobacteria
2	31	86.1	207	1 KGUA_ECO57	Q8x388 escherichia
3	30	83.3	157	1 YH13_STRPN	P58260 streptococc
4	30	83.3	207	1 KGUA_ECOLI	P24234 escherichia
5	30	83.3	207	1 KGUA_SALTI	Q8x2h9 salmonella
6	30	83.3	207	1 KGUA_SALTY	Q9x6m5 salmonella
7	30	83.3	207	1 KGUA_YERPE	Q8z1q2 yersinia pe
8	30	83.3	576	1 DNAB_MYCIT	Q9f5p4 mycobacteri
9	29	80.6	202	1 UMPK_ARATH	Q04905 arabidopsis
10	29	80.6	312	1 TF2B_ORYSA	Q8w0w3 oryza sativ
11	29	80.6	517	1 LEGA_PEA	P02857 pisum sativ
12	29	80.6	520	1 LEG2_PEA	P15838 pisum sativ
13	29	80.6	674	1 COAA_BACTJ	O87905 bacillus th
14	28	77.8	126	1 CV02_HUMAN	Q9y3m2 homo sapien
15	28	77.8	127	1 CV02_MOUSE	Q9d1c2 mus musculu
16	28	77.8	178	1 YK01_ARCFU	O28278 archaeoglob
17	28	77.8	199	1 KGUA_ANASP	Q8z017 anabaena sp
18	28	77.8	242	1 RNPH_NEIMA	Q9jtn2 neisseria m
19	28	77.8	242	1 RNPH_NEIMB	Q9jyn5 neisseria m
20	28	77.8	260	1 DCK_HUMAN	P27707 homo sapien
21	28	77.8	260	1 DCK_MOUSE	P43346 mus musculu
22	28	77.8	386	1 NESG_HUMAN	Q9ul16 homo sapien
23	28	77.8	395	1 PELO_DROME	P48612 drosophila
24	28	77.8	445	1 MEK1_SCHPO	Q10292 schizosacch
25	28	77.8	476	1 YHGF_NEIGO	Q51062 neisseria g
26	28	77.8	509	1 AGP4_MOUSE	Q9whv6 mus musculu
27	28	77.8	517	1 Y967_TREPA	O83933 treponema p
28	28	77.8	676	1 HMUR_YERPE	Q56989 yersinia pe
29	28	77.8	687	1 HMUR_YERN	P31493 yersinia en
30	28	77.8	757	1 YHGF_NEIMA	P57092 neisseria m
31	28	77.8	757	1 YHGF_NEIMB	Q51152 neisseria m
32	28	77.8	917	1 SMOO_HUMAN	P53814 homo sapien
33	28	77.8	1038	1 CIN8_YEAST	P27895 saccharomyc

34	28	77.8	1184	1 B1MC_EMENI	P17120 emericella
35	28	77.8	1714	1 YP93_CAEEL	Q09475 caenorhabdi
36	28	77.8	1816	1 LMA4_MOUSE	P97927 mus musculu
37	27	75.0	133	1 TOX5_BORPE	P04981 bordetella
38	27	75.0	204	1 VNSC_SENDF	P14253 sendai viru
39	27	75.0	204	1 VNSC_SENDF	P14253 sendai viru
40	27	75.0	204	1 VNSC_SENDF	P14253 sendai viru
41	27	75.0	204	1 VNSC_SENDF	P14253 sendai viru
42	27	75.0	209	1 SEG8_BPT4	P19895 bacterioph
43	27	75.0	235	1 GAG_MLVAB	P03333 abelson mur
44	27	75.0	239	1 RNPH_RICCN	Q92977 rickettsia
45	27	75.0	240	1 RNPH_RICPR	Q92977 rickettsia

ALIGNMENTS

RESULT 1
KGUA_FUSNN
ID KGUA_FUSNN STANDARD; PRT: 185 AA.
AC Q8RH19;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMP OR FN2033
OS Fusobacterium nucleatum (subsp. nucleatum)
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21883394; PubMed=11899109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner M., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Fusch G., Haselkorn R.,
RA Ponstein M., Kypides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586".
RT J. Bacteriol. 184:2005-2018(2002).
RL -!- FUNCTION: Essential for recycling GMP and indirectly, CGMP.
CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE010506; AAL94118.1; --
CC HAMAP: MF_00328; -- 1.
CC InterPro: IPR000619; Guanylate_kin.
CC Pfam: PF00625; Guanylate_kin; 1.
CC SMART: SM00072; GuK; 1.
CC PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE: PS00952; GUANYLATE_KINASE_2; 1.
CC Transferrase; Kinase; ATP-binding; Complete proteome.
KW NP_BIND 11 18
FT ATP (BY SIMILARITY).
SQ SEQUENCE 185 AA; 21025 MW; 988EF86443662B96 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 185;
Best Local Similarity 100.0%; Pred.No. 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6

DB 131 RLGRNQ 136

RESULT 2	KGUA_ECO57	STANDARD;	PRT;	207 AA.
ID	KGUA_ECO57	STANDARD;	PRT;	207 AA.
AC	Q8XD88;			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Guanylate kinase (EC 2.7.4.8) (GMP kinase).			
GN	GMP OR 25074 OR EC34523.			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OC	NCBI_TaxID=83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RC	MEDLINE=21074935; PubMed=11206551;			
RA	Ferra N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,			
RA	Rose D.J., Maynew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamatzis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Welch R.A., Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"			
RL	Nature 409:529-533(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / RIMD 0509952;			
RC	MEDLINE=21156231; PubMed=11258796;			
RA	Hayashi T., Makino K., Okuyama K., Kurokawa K., Ishii K., Yokoyama K.,			
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,			
RA	Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,			
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.;			
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli			
RT	O157:H7 and genomic comparison with a laboratory strain K-12.;"			
RL	DNA Res. 8:11-22(2001).			
CC	-1- FUNCTION: Essential for recycling GMP and indirectly, cGMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
EMBL	AE005592; AAC58792.1; -			
EMBL	AP002566; BAB37946.1; -			
PIR	C91194; C91194			
PIR	D86041; D86041			
DR	HAWAP; MF_003287; - 1.			
DR	InterPro; IPR000619; Guanylate_kin.			
DR	Pfam; PF00625; Guanylate_kin; 1.			
DR	SMART; SM00072; GuK; 1.			
DR	PROSITE; PS00856; GUANYLATE KINASE 1; 1.			
DR	PROSITE; PS00052; GUANYLATE KINASE 2; 1.			
KW	Transferase; Kinase; ATP-binding; Complete proteome.			
FT	NP BIND 11 18 ATP (BY SIMILARITY).			
SQ	SEQUENCE 207 AA; 23651 MW; CE669DB404E481E6 CRC64;			
Query Match	86.1%; Score 31; DB 1; Length 207;			
Best Local Similarity	85.7%; Pred. No. 6.7;			
Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Oy	1 RLGRNQ 7			
DB	134 RLGRDQ 140			

94 KLRGRNE 100

RESULT 4

D_KGUA_ECOLI STANDARD; PRT; 207 AA.

C P24234; 1992 (Rel. 21, Created)

T 01-MAR-1992 (Rel. 21, Last sequence update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

N Guanylate kinase (EC 2.7.4.8) (GMP kinase).

E GMP OR SPOR OR B3648 OR C4473 OR SF3688.

N Escherichia coli.

S Escherichia coli O6, and

S Shigella flexneri.

S Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

C Enterobacteriaceae; Escherichia.

N CBI_TaxID=562, 217992, 623;

N [1]

N SEQUENCE FROM N.A.

P SPECIES=E.coli; STRAIN=K12;

C MEDLINE=93300828; PubMed=8390989;

X Gentry D., Bengra C, Ikebara K., Caehel M.;

N "Guanylate Kinase of Escherichia coli K-12.";

T J. Biol. Chem. 268:14316-14321(1993).

N [2]

N SEQUENCE FROM N.A.

C SPECIES=E.coli; STRAIN=K12 / MG1655;

C MEDLINE=93315143; PubMed=7686882;

X Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;

N "DNA sequence and analysis of 136 kilobases of the Escherichia coli

T genome: organizational symmetry around the origin of replication.";

T Genomics 16:551-561(1993).

N [3]

N SEQUENCE FROM N.A.

C SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;

X MEDLINE=22388234; PubMed=12471157;

C Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

X Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

X Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

X Mobley H.L.T., Sonnenberg M.S., Blattner F.R.;

N "Extensive mosaic structure revealed by the complete genome sequence

T of uropathogenic Escherichia coli.";

T Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

N [4]

N SEQUENCE FROM N.A.

C SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;

X MEDLINE=22272406; PubMed=12384590;

C Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

X Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,

X Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

X Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,

X Yu J.;

N "Genome sequence of Shigella flexneri 2a: insights into pathogenicity

T through comparison with genomes of Escherichia coli K12 and O157.";

T Nucleic Acids Res. 30:4432-4441(2002)

N C -!- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.

C C -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.

C C -!- SUBUNIT: HOMOTETRAMER (UNDER LOW IONIC CONDITIONS) OR HOMODIMER

C C (UNDER HIGH IONIC CONDITIONS).

C C -!- SUBCELLULAR LOCATION: Cytoplasmic.

C C -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.

C C

C This SWISS-PROT entry is copyright. It is produced through a collaboration

C between the Swiss Institute of Bioinformatics and the EMBL outstation -

C the European Bioinformatics Institute. There are no restrictions on its

C use by non-profit institutions as long as its content is in no way

C modified and this statement is not removed. Usage by and for commercial

C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

C or send an email to license@isb-sib.ch.

C

C

DR EMBL; M84400; AAB88711.1; --

AA02003

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL627280; CAD03253.1; -
CC EMBL; AE016846; AAO71261.1; -
CC DR
CC HAMAP; MF_00328; -; 1.
CC DR
CC InterPro; IPR000619; Guanylate_kin.
CC PFam; PF00625; Guanylate_kin; 1.
CC DR
CC SMART; SM00072; GuKc; 1.
CC DR
CC PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
CC DR
CC PROSITE; PS00052; GUANYLATE_KINASE_2; 1.
CC DR
CC TRANSFERASE; Kinase; ATP-binding; Complete proteome.
CC NP_BIND 11 18 ATP (BY SIMILARITY).
CC FT
CC SEQUENCE 207 AA; 23498 MW; A26757C1FB40970B CRC64;
CC
CC Query Match 83.3%; Score 30; DB 1; Length 207;
CC Best Local Similarity 85.7%; Pred. No. 11;
CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 RLGRNQ 7
CC DB 134 RLGRGQ 140
CC
CC RESULT 6
CC KGUA_SALTY
CC ID_KGUA_SALTY STANDARD; PRT; 207 AA.
CC AC QX6M5;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
CC GN GMK OR STM3740.
CC OS Salmonella typhimurium.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Salmonella.
CC OC NCBI_TaxID=602;
CC FN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=LT2;
CC RA Beck B.J.; Huelsmeyer M.; Downs D.M.;
CC RT "Salmonella typhimurium guanylate kinase."
CC RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
CC RX MEDLINE=21534948; PubMed=11677609;
CC RA McClelland M.; Sanderson K.E.; Spieth J.; Clifton S.W.; Latreille P.;
CC RA Courtney L.; Porwollik S.; Ali J.; Dante M.; Du F.; Hou S.; Layman D.;
CC RA Leonard S.; Nguyen C.; Scott K.; Holmes A.; Grewal N.; Mulvaney E.;
CC RA Ryan E.; Sun H.; Florea L.; Miller W.; Stoneking T.; Nhan M.;
CC RA Waterston R.; Wilson R.K.;
CC RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
CC LT2."
CC RL Nature 413:852-856 (2001).
CC CC -!- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.
CC CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC CC -!- SUBUNIT: HOMOTETRAMER (UNDER LOW IONIC CONDITIONS) OR HOMODIMER
CC CC (UNDER HIGH IONIC CONDITIONS) (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF140283; AAD31506.1; -
CC EMBL; AE008874; AAL22599.1; -
CC HSP; F15454; IGKI.
CC StyGene; SG77777; gmK.

```

```

DR HAMAP; MF_00328; -; 1.
DR InterPro; IPR000619; Guanylate_kin.
DR PFam; PF00625; Guanylate_kin; 1.
DR SMART; SM00072; GuKc; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS00052; GUANYLATE_KINASE_2; 1.
DR TRANSFERASE; Kinase; ATP-binding; Complete proteome.
DR NP_BIND 11 18 ATP (BY SIMILARITY).
DR FT
DR SEQUENCE 207 AA; 23498 MW; A26757C1FB40970B CRC64;
DR
DR Query Match 83.3%; Score 30; DB 1; Length 207;
DR Best Local Similarity 85.7%; Pred. No. 11;
DR Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DR
DR QY 1 RLGRNQ 7
DR DB 134 RLGRGQ 140
DR
DR RESULT 7
DR KGUA_YERPE
DR ID_KGUA_YERPE STANDARD; PRT; 207 AA.
DR AC Q8ZJQ2;
DR DT 28-FEB-2003 (Rel. 41, Created)
DR DT 28-FEB-2003 (Rel. 41, Last sequence update)
DR DT 28-FEB-2003 (Rel. 41, Last annotation update)
DR DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
DR GN GMK OR YPO0040 OR Y0101.
DR OS Yersinia pestis
DR OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
DR OC Enterobacteriaceae; Yersinia.
DR OC NCBI_TaxID=632;
DR FN [1]
DR RP SEQUENCE FROM N.A.
DR RC STRAIN=CO-92 / Biovar Orientalis;
DR RX MEDLINE=21470413; PubMed=11586360;
DR RA Parkhill J.; Wren B.W.; Thomson N.R.; Titball R.W.; Holden M.T.G.;
DR RA Prentice M.B.; Sebahia M.; James K.D.; Churcher C.; Mungall K.L.;
DR RA Baker S.; Basham D.; Bentley S.D.; Brooks K.; Cerdano-Tarraga A.M.;
DR RA Chillingworth T.; Cronin A.; Davies R.M.; Davis P.; Dougan G.;
DR RA Feltwell T.; Hamlin N.; Holroyd S.; Jags K.; Karlyshev A.V.;
DR RA Leather S.; Moule S.; Oyston P.C.F.; Quail M.; Rutherford K.;
DR RA Simmonds M.; Skelton J.; Stevens K.; Whitehead S.; Barrell B.G.;
DR RT "Genome sequence of Yersinia pestis, the causative agent of plague."
DR RL Nature 413:523-527 (2001).
DR RN [2]
DR RP SEQUENCE FROM N.A.
DR RC STRAIN=KIM5 / Biovar Mediaevalis;
DR RX MEDLINE=22137863; PubMed=12142430;
DR RA Deng W.; Burland V.; Plunkett G. III; Boutin A.; Mayhew G.F.; Liss P.;
DR RA Perna N.T.; Rose D.J.; Mau B.; Zhou S.; Schwartz D.C.;
DR RA Fetherston J.D.; Lindler L.E.; Brubaker R.R.; Plano G.V.;
DR RA Straley S.C.; McDonough K.A.; Nilles M.L.; Matson J.S.; Blattner F.R.;
DR RA Perry R.D.;
DR RT "Genome sequence of Yersinia pestis KIM."
DR RL J. Bacteriol. 184:4601-4611 (2002).
DR CC -!- FUNCTION: Essential for recycling GMP and indirectly, CGMP.
DR CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
DR CC -!- SUBUNIT: HOMOTETRAMER (UNDER LOW IONIC CONDITIONS) OR HOMODIMER
DR CC CC (UNDER HIGH IONIC CONDITIONS) (By similarity).
DR CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.
DR
DR This SWISS-PROT entry is copyright. It is produced through a collaboration
DR between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR the European Bioinformatics Institute. There are no restrictions on its
DR use by non-profit institutions as long as its content is in no way
DR modified and this statement is not removed. Usage by and for commercial
DR entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR or send an email to license@isb-sib.ch).
DR
DR EMBL; AJ414141; CAC88907.1; -
DR EMBL; AE013610; AAM83695.1; -
DR PIR; A10005; A10005.
DR HAMAP; MF_00328; -; 1.

```


DR InterPro; IPR000619; Guanylate kin.
 DR Pfam; PF00625; Guanylate_kin; 1.
 DR SMART; SMC0072; GuKc; 1.
 DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE; PS00856; GUANYLATE_KINASE_2; 1.
 DR TRANSFERASE; Kinase; ATP-binding; Complete proteome.
 KW NP_BIND 11 18 ATP SIMILARITY)
 FT NP_BIND 11 18
 SQ SEQUENCE 207 AA; 23503 MW; 01FDCD618EFED91E CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 207;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RLGRNQ 7
 Db 134 RLGRGQ 140
 RESULT 8
 DNAB_MYCIT STANDARD; PRT; 576 AA.
 ID DNAB_MYCIT
 AC Q9F5P4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Replicative DNA helicase (EC 3.6.1.-) [Contains: Min dnaB intein]
 DE (Fragment).
 GN DNAB.
 OS Mycobacterium intracellulare.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1442;
 RA Madiraju M.V.V.S., Yamamoto K., Rajagopalan M., Rutherford S.A.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PARTICIPATES IN INITIATION AND ELONGATION DURING
 CC CHROMOSOME REPLICATION; IT EXHIBITS DNA-DEPENDENT ATPASE ACTIVITY
 CC (BY SIMILARITY).
 CC -1- PWM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. DNAB SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF307984; AAC31144.1;
 DR InterPro; IPR001198; DnaB_helicase.
 DR InterPro; IPR006141; INTEIN.
 DR InterPro; IPR006142; INTEIN.
 DR InterPro; IPR004042; Intein_endonuc.
 DR Pfam; PF03796; DnaB_C; 1.
 DR PRINTS; PR00379; INTEIN.
 DR ProDom; PD332834; DnaB_C; 1.
 DR TIGRFAMs; TIGR01443; Intein_Cterm; 1.
 DR PROSITE; PS00818; INTEIN_C_TER; FALSE_NEG.
 DR PROSITE; PS00819; INTEIN_ENDONUCLEASE; 1.
 DR PROSITE; PS00817; INTEIN_N_TER; FALSE_NEG.
 DR DNA replication; DNA-binding; ATP-binding; Hydrolase; Helicase;
 KW Primosome; Autocatalytic cleavage; Protein splicing.
 FT NON_TER 1 15
 FT CHAIN <1 15
 FT REPLICATIVE DNA HELICASE, 1ST PART
 FT (POTENTIAL).
 FT CHAIN 16 351
 FT MIN DNAB INTEIN (POTENTIAL).
 FT CHAIN 352 576
 FT REPLICATIVE DNA HELICASE, 2ND PART
 FT (POTENTIAL).
 FT

SQ SEQUENCE 576 AA; 63529 MW; 6D59341A779D082F CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 576;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RLGRNQ 7
 Db 268 RVGRNE 274
 RESULT 9
 UMPK_ARATH STANDARD; PRT; 202 AA.
 ID UMPK_ARATH
 AC O04905;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Uridylate kinase (EC 2.7.4.-) (UK) (Uridine monophosphate kinase)
 DE (UMP kinase) (UMP/CMP kinase).
 GN PYR6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=98245083; PubMed=9576794;
 RA Zhou L., Lacroute F., Thornburg R.;
 RT "Cloning, expression in Escherichia coli, and characterization of
 RT Arabidopsis thaliana UMP/CMP kinase.";
 RL Plant Physiol. 117:245-254(1998).
 CC -1- FUNCTION: FURNISHES THE CELL WITH PYRIMIDINES. ACCEPTS UMP AND
 CC CMP AS PHOSPHATE ACCEPTORS WITH HIGH ACTIVITY; DOES NOT ACT ON
 CC DMP AND DUMP.
 CC -1- CATALYTIC ACTIVITY: ATP + UMP = ADP + UDP.
 CC -1- CATALYTIC ACTIVITY: ATP + CMP = ADP + CDP.
 CC -1- SIMILARITY: Belongs to the adenylylate kinase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF000147; AAB71135.1;
 DR HSPF; P20425; 3URD.
 DR InterPro; IPR000850; Adenylylate_kin.
 DR InterPro; IPR006266; UMP_CMP_kinase.
 DR Pfam; PF00406; ADK; 1.
 DR PRINTS; PR00094; ADENYLTKINASE.
 DR ProDom; PD000657; Adenylylate_kin; 1.
 DR TIGRFAMs; TIGR01359; UMP_CMP_kin_fam; 1.
 DR PROSITE; PS00113; ADENYLATE_KINASE; 1.
 DR TRANSFERASE; Kinase; Pyrimidine biosynthesis; ATP-binding.
 KW NP_BIND 21 29
 FT NP_BIND 21 29
 FT ATP.
 SQ SEQUENCE 202 AA; 22482 MW; 41AD0CFACD816315 CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 202;
 Best Local Similarity 85.7%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RLGRNQ 7
 Db 133 RLGRNQ 139
 RESULT 10
 TF2B_ORYSA STANDARD; PRT; 312 AA.
 ID TF2B_ORYSA

AC Q8W03;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription initiation factor IIB (General transcription factor
 DE TFIIIB).
 GN TFIIIB.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzoideae; Oryza.
 CC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21957784; PubMed=11971135;
 RA Zhu Q., Ordiz M.I., Dabi T., Beachy R.N., Lamb C.;
 RT "Rice TATA binding protein interacts functionally with transcription
 RT factor IIB and the RFX2 bZIP transcriptional activator in an enhanced
 RT plant in vitro transcription system.";
 RL Plant Cell 14:795-803(2002).
 CC -1- FUNCTION: GENERAL FACTOR THAT PLAYS A MAJOR ROLE IN THE ACTIVATION
 CC OF EUKARYOTIC GENES TRANSCRIBED BY RNA POLYMERASE II.
 CC -1- COFACTOR: Binds a zinc ion per subunit (by similarity).
 CC -1- SUBUNIT: ASSOCIATES WITH TFIIID-11A (DA COMPLEX) TO FORM TFIIID-
 CC 11A-IIB (DAB-COMPLEX) WHICH IS THEN RECOGNIZED BY POLYMERASE II.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the TFIIIB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 DR EMBL; AF464908; AAL73491.1; -;
 DR Gramene; Q8W03; -;
 DR InterPro; IPR006670; Cyclin.
 DR Pfam; PF00382; transcript_fac2; 2.
 DR PRINTS; PR00685; TIFACTOR.IIB.
 DR SMART; SM00385; CYCLIN; 2.
 DR PROSITE; PS00782; TFIIIB; 1.
 KW Transcription regulation; Nuclear protein; Repeat; Zinc-finger;
 KW Metal-binding; Zinc. 29 ZN-RIBBON TFIIIB-TYPE.
 FT ZN-FING 115 192
 FT REPEAT 216 290
 FT METAL 6 6
 FT METAL 9 9
 FT METAL 26 26
 FT METAL 29 29
 FT METAL 312 312
 SQ SEQUENCE 312 AA; 34298 MW; 338435F8048C46AF CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 312;
 Best Local Similarity 83.3%; Pred. No. 30;
 Matches .5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LRGRNQ 7
 DB 142 IRRGNQ 147
 RESULT 11
 LEGA_PEA
 ID LEGA_PEA STANDARD; PRT; 517 AA.
 AC P02857; Q41033; Q41034;
 DT 21-JUL-1986 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Legumin A precursor.
 GN LEGA.

OS Pisum sativum (Garden pea).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 CC NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Feltham First;
 RA MEDLINE=84247316; PubMed=6330672;
 RA Lycett G.W., Croy R.R.D., Shirat A.H., Boulter D.;
 RT "The complete nucleotide sequence of a legumin gene from pea (Pisum
 RT sativum L.).";
 RL Nucleic Acids Res. 12:4493-4506(1984).
 RN [2]
 RP SEQUENCE OF 209-411 FROM N.A.
 RA Domoney C., Barker D., Casey R.;
 RT "The complete deduced amino acid sequences of legumin beta-
 RT polypeptides from different genetic loci in Pisum.";
 RL Plant Mol. Biol. 7:467-474(1986).
 CC -1- FUNCTION: THIS PROTEIN FOUND IN THE SEEDS OF MANY LEGUMINOUS &
 CC NONLEGUMINOUS PLANTS IS THE SOURCE OF SULFUR-CONTAINING AMINO
 CC ACIDS IN SEED MEALS.
 CC -1- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
 CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: Belongs to the IIS seed storage protein (globulins)
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 DR EMBL; X02982; CAA26720.1; ALT_SEQ.
 DR EMBL; M16904; AAA33677.1; -;
 DR PIR; A22866; FWPMLA.
 DR PIR; S09559; S09559.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin_sup_11s.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin; 2.
 DR PRINTS; PR00439; IISGLOBULIN.
 DR PROSITE; PS00305; IIS SEED STORAGE; 1.
 KW Seed storage protein; Multigene family; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 332 ALPHA CHAIN (ACIDIC).
 FT CHAIN 333 517 BETA CHAIN (BASIC).
 FT DISULFID 107 339 INTERCHAIN (ALPHA-BETA) (POTENTIAL).
 FT DOMAIN 89 122 GLN/GLU-RICH.
 FT DOMAIN 189 208 GLN-RICH.
 FT DOMAIN 260 313 ARG/GLN/GLU-RICH.
 FT CONFLICT 219 219 Y -> F (IN REF. 2).
 FT CONFLICT 319 319 G -> R (IN REF. 2).
 FT SEQUENCE 517 AA; 58805 MW; CD2CA2B41758D0BA CRC64;
 SQ
 Query Match 80.6%; Score 29; DB 1; Length 517;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RLGRNQ 7
 DB 233 RLQGRNE 239
 RESULT 12
 LEG2_PEA
 ID LEG2_PEA STANDARD; PRT; 520 AA.
 AC P15958;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Legumin A2 precursor.
 GN LEGA2.
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Greenfeast;
 RX MEDLINE=90174993; PubMed=2038850;
 RA Rerie W.G., Whitecross M.I., Higgins T.J.V.;
 RT "Nucleotide sequence of an A-type legumin gene from pea."
 RL Nucleic Acids Res. 18:655-655(1990).
 CC -!- FUNCTION: THIS PROTEIN FOUND IN THE SEEDS OF MANY LEGUMINOUS &
 CC NONLEGUMINOUS PLANTS IS THE SOURCE OF SULFUR-CONTAINING AMINO
 CC ACIDS IN SEED MEALS.
 CC -!- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
 CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
 CC DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X17193; CAA35056.1; --
 CC PIR; S08237; S08237.
 CC InterPro; IPR006045; Cupin.
 CC InterPro; IPR007113; Cupin_sup.
 CC InterPro; IPR006044; Seedstore_11s.
 CC Pfam; PF00190; Cupin; 2.
 CC PRINTS; PR00439; 11S_GLOBULIN.
 CC PROSITE; PS00305; 11S_SEED_STORAGE; 1.
 CC Seed storage protein; Multigene family; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 335 ALPHA CHAIN (ACIDIC).
 FT CHAIN 336 520 BETA CHAIN (BASIC).
 FT DISULFID 108 342 INTERCHAIN (ALPHA-BETA) (POTENTIAL).
 SQ SEQUENCE 520 AA; 59269 MW; CDB8B5E350C760C6 CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 520;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RLGRNQ 7
 D5 234 RLQGRNE 240
 RESULT 13
 ID CORA_BACTU STANDARD; PRT; 674 AA.
 AC C87905;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peptidicidal crystal protein cry24Aa (Insecticidal delta-endotoxin
 DE CryXXIVA(a)) (Crystalline entomocidal protoxin) (Crystal protein)
 DE (Insecticidal protein Jegg72) (Fragment).
 OS Bacillus thuringiensis (subsp. jegathesan).
 GN Cry24AA OR CRYXXIVA(A).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=56955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawalek M.D., Gill S.S.;

"Isolation and characterization of insecticidal genes from *Bacillus thuringiensis* subsp. *jegathesan*.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF INSECTS.
 CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOREFORMATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -!- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -!- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U88188; AAC61891.1; --
 CC InterPro; IPR001178; Endotoxin.
 CC InterPro; IPR005638; endotoxin_C.
 CC InterPro; IPR005639; endotoxin_N.
 CC Pfam; PF00555; endotoxin; 1.
 CC Pfam; PF03944; endotoxin_C; 1.
 CC Pfam; PF03945; endotoxin_N; 1.
 CC Toxin; Sporulation.
 KW NON_TER 674 674
 FT SEQUENCE 674 AA; 75959 MW; DA3904DAB891C978 CRC64;
 SQ
 Query Match 80.6%; Score 29; DB 1; Length 674;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LRGRNQ 7
 D5 435 LRGRNQ 440
 RESULT 14
 ID CV02_HUMAN STANDARD; PRT; 126 AA.
 AC Q9Y3M2; OSUK9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein C22orf2 (Cytosolic leucine-rich protein) (HRIHFB2025).
 GN C22ORF2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wienann S., Weil B., Wellenreuther R., Gassenhuber J., Glassel S.,
 RA Ansgorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesternoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Rampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Collins J.E., Huckle E.J.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,

RA Ciamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Beggley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhamsi P.D., Dockree C., Dodingworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tanlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay C.N., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Ocell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.B., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Sprigdon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuoyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaul S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton D., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing J.,
RA Schett P., Walker C., Wamley A., Wohldmann P., Pepin K., Nelson J.,
RA Korfi I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Sudar M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Serousi E., Fransson I., Tapia I., Bruder C.B., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiahun Y., Wright H.,
RA "The DNA sequence of human chromosome 22."
RA Nature 402:489-495(1999).
RA [4]
RA SEQUENCE FROM N.A.
RA Huang C.-H.;
RA "A novel cytosolic leucine-rich protein."
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RA [5]
RA SEQUENCE FROM N.A.
RA TISSUE=Skin;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman N., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RA [6]
RA SEQUENCE OF 60-126 FROM N.A.

TISSUE=Fetal brain;
RA Ueki N.;
RA "HRI NTF human fetal brain cDNA project."
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RA -1- TISSUE SPECIFICITY: Found in heart, brain, lung, liver, muscle,
RA kidney and testis.
RA -----
RA This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (see http://www.isb-sib.ch/announce/
RA or send an email to license@isb-sib.ch).
RA -----
RA EMBL; AL136686; CAB66621.1; -
RA EMBL; AL050345; CAB43547.1; -
RA EMBL; AL021707; CAB363442.1; -
RA EMBL; AF331041; AAL56062.1; -
RA EMBL; BC016139; AAH16139.1; -
RA EMBL; AB015347; BA888119.1; -
RA EMBL; HGNC:1307; C22orf2. -
RA GENBANK; 77 98 LEUCINE-ZIPPER.
RA DOMAIN 73 73 E -> K (IN REF. 4).
RA FT CONFLICT 73 73
RA SQ SEQUENCE 126 AA; 14470 MW; 0D243AD2CC436E55 CRC64;
RA
RA Query Match 77.8%; Score 28; DB 1; Length 126;
RA Best Local Similarity 85.7%; Pred. No. 19;
RA Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RA
RA QY 1 RLGRNQ 7
RA |||||
RA Db 76 RLRRNQ 82
RA
RA RESULT 15
RA CV02 MOUSE
RA ID CV02 MOUSE STANDARD; PRT; 127 AA.
RA AC Q9DIC2;
RA DT 28-FEB-2003 (Rel. 41, Last sequence update)
RA DT 28-FEB-2003 (Rel. 41, Last sequence update)
RA DT 15-SEP-2003 (Rel. 42, Last annotation update)
RA DE Protein C22orf2 homolog (Cytosolic leucine-rich protein).
RA C22ORF2.
RA GN Mus musculus (Mouse).
RA OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RA OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RA OX NCBI_TaxID=10090;
RA RN [1]
RA RP SEQUENCE FROM N.A.
RA Huang C.-H.;
RA "A novel cytosolic leucine-rich protein."
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RA [2]
RA SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Embryo;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli R., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

```

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- TISSUE SPECIFICITY: Found in heart, brain, lung, liver, muscle,
CC kidney and testis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AF331040; AAL56061.1; -
CC DR EMBL; AK003719; BAB22956.1; -
CC DR EMBL; BC005733; AAO05733.1; -
CC FT DOMAIN 77 98 LEUCINE-ZIPPER.
CC SEQUENCE 127 AA; 14534 MW; 7CE38AA48D045CCC7 CRC64;
CC -----
Query Match 77.8%; Score 28; DB 1; Length 127;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
DB 76 RLGRNQ 82

```

Search completed: February 11, 2004, 17:04:12
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-8

Perfect score: 36

Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	31	86.1	75	2 AB1193	hypothetical prote
2	31	86.1	145	2 AD2345	hypothetical prote
3	31	86.1	195	2 P36734	probable transcrip
4	31	86.1	207	2 C31134	guanylate kinase [
5	31	86.1	207	2 D86041	guanylate kinase [
6	31	86.1	213	1 S71287	myb-related 24.7K
7	31	86.1	280	2 AF1375	B. subtilis ynfP
8	31	86.1	281	2 AD1745	B. subtilis ynfP
9	31	86.1	386	2 AB1996	hypothetical prote
10	30	83.3	157	2 P35139	conserved hypotet
11	30	83.3	177	2 D98066	conserved hypotet
12	30	83.3	207	1 KIECGU	guanylate kinase (
13	30	83.3	207	2 A10005	guanylate kinase (
14	30	83.3	207	2 A10969	5'guanylate kinase
15	30	83.3	445	2 T21262	hypothetical prote
16	30	83.3	1050	2 T36395	hypothetical prote
17	29	80.6	170	2 AB2402	hypothetical prote
18	29	80.6	184	2 B42707	transposase homolo
19	29	80.6	203	2 S09559	legumin A (clone p
20	29	80.6	203	2 F72273	conserved hypotet
21	29	80.6	251	2 AB0213	conserved hypotet
22	29	80.6	307	2 PC2050	legumin type A alp
23	29	80.6	315	2 JC2094	legumin type A alp
24	29	80.6	482	2 S49877	legumin A precursor
25	29	80.6	497	2 S14392	legumin A1 precurs
26	29	80.6	498	2 S44294	legumin A4 precurs
27	29	80.6	500	2 S14393	legumin A4 precurs
28	29	80.6	507	2 T06452	probable legumin A
29	29	80.6	517	1 FWPMLA	legumin A precursor

30	29	80.6	520	2 S08237	legumin A2 precurs
31	29	80.6	779	2 S36111	osteoblast-specifi
32	29	80.6	811	2 S36109	osteoblast-specifi
33	29	80.6	836	2 S36110	osteoblast-specifi
34	29	80.6	2610	2 T20968	hypothetical prote
35	28	77.8	75	2 AH1550	hypothetical prote
36	28	77.8	123	2 D84161	hypothetical prote
37	28	77.8	144	2 T49706	hypothetical prote
38	28	77.8	178	2 H69499	conserved hypotet
39	28	77.8	193	2 E83864	thiorredoxin BH1717
40	28	77.8	199	2 AB1820	guanylate kinase [
41	28	77.8	242	2 AB1856	tRNA nucleotidyltr
42	28	77.8	242	2 G81076	ribonuclease PH NM
43	28	77.8	260	1 A38585	deoxycytidine kina
44	28	77.8	260	2 A55122	deoxycytidine kina
45	28	77.8	278	2 AD2325	hypothetical prote

ALIGNMENTS

RESULT 1

AB1193

hypothetical protein lmo0946 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1193

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, F.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1193

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <GLA>

A:Cross-references: GB:NC 003210; PIDN:CAC99024.1; PID:g16410348; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo0946

Query Match 86.1%; Score 31; DB 2; Length 75;
Best Local Similarity 100.0%; Pred.No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6
DB 70 RLGRNQ 75

RESULT 2

AD2345

hypothetical protein all4315 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AD2345

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2345

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076014.1; PID:gl7133451; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4315.

Query Match 86.1%; Score 31; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
 |||||
 99 RLGRN 104

Db

RESULT 3
 F96734
 Probable transcription factor F2N20.2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: F96734
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Hughes, B.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96734
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-195 <STO>
 A:Cross-references: GB:AE005173; NID:g6714312; PIDN:AAF26005.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F2N20.2
 A:Map position: 1
 C:Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology

Query Match 86.1%; Score 31; DB 2; Length 195;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
 |||||
 22 RLGRN 27

Db

RESULT 4
 C91194
 Guanylate kinase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: C91194
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: C91194
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-207 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA37946.1; PID:g13363998; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: EC84523
 C:Superfamily: guanylate kinase; guanylate kinase homology

Query Match 86.1%; Score 31; DB 2; Length 207;
 Best Local Similarity 85.7%; Pred. No. 25;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 7
 |||||
 1 RLGRN 7

Db

RESULT 5
 D86041
 guanylate kinase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: D86041
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: D86041
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-207 <STO>
 A:Cross-references: GB:AE005174; NID:g12518407; PIDN:AAG58792.1; GSPDB:GN00145; UWGP:2E
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gmk
 C:Superfamily: guanylate kinase; guanylate kinase homology

Query Match 86.1%; Score 31; DB 2; Length 207;
 Best Local Similarity 85.7%; Pred. No. 25;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 7
 |||||
 134 RLGRDQ 140

Db

RESULT 6
 S71287
 myb-related 24.7K transcription factor - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 16-Jun-2000
 C:Accession: S71287; JC5729
 R:Kirik, V.; Baumlein, H.
 submitted to the EMBL Data Library, December 1995
 A>Description: Isolation and chromosomal localization of leaf-specific cDNA encoding fc
 A:Reference number: S71287
 A:Accession: S71287
 A:Molecule type: mRNA
 A:Residues: 1-213 <KIR>
 A:Cross-references: EMBL:Z68157; NID:g1197189; PIDN:CAA92280.1; PID:g1197190
 R:Kirik, V.; Baumlein, H.
 Gene 183, 109-113, 1996
 A>Title: A novel leaf-specific myb-related protein with a single binding repeat.
 A:Reference number: JC5729; MUID:97149286; PMID:8996094
 A:Accession: JC5729
 A:Molecule type: mRNA
 A:Residues: 1-213 <K12>
 A:Cross-references: EMBL:Z68157; NID:g1197189; PIDN:CAA92280.1; PID:g1197190
 A:Comment: This protein is involved in the control of anthocyanin biosynthesis and the
 C:Genetics:
 A:Gene: Atmyb12
 A:Map position: 1
 C:Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology

Query Match 86.1%; Score 31; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
 |||||
 22 RLGRN 27

Db

RESULT 7

AF13175
B. subtilis YnfF protein homolog lmc2406 [imported] - Listeria monocytogenes (strain EGD)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF13175
R.; Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Sincos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A.; Titis: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF13175
A:Molecule type: DNA
A:Residues: 1-280 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00484.1; PID:g16411894; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmc2406

Query Match 86.1%; Score 31; DB 2; Length 280;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
Db 192 RLGRNQ 199

RESULT 8

AD1745
B. subtilis YnfF protein homolog lin2505 [imported] - Listeria innocua (strain Clp11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD1745
R.; Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Sincos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A.; Titis: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1745
A:Molecule type: DNA
A:Residues: 1-281 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97732.1; PID:g16415027; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2505

Query Match 86.1%; Score 31; DB 2; Length 281;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
Db 193 RLGRNQ 199

RESULT 9

AE1996
hypothetical protein all1523 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AE1996
R.; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB1807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA577889.1; PID:g17135343; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1523

Query Match 86.1%; Score 31; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLGRNQ 6
Db 299 RLGRNQ 304

RESULT 10

F95199
conserved hypothetical protein SPI713 [imported] - Streptococcus pneumoniae (strain TIC)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: F95199
R.; Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heion, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrisior
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95199
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <KUR>
A:Cross-references: GB:AB005672; PIDN:AAK75791.1; PID:g14973208; GSPDB:GN00164; TIGR:SE
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI713
C:Superfamily: conserved hypothetical protein HI0943

Query Match 83.3%; Score 30; DB 2; Length 157;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
Db 94 RLGRNQ 100

RESULT 11

D98066
conserved hypothetical protein spr1557 [imported] - Streptococcus pneumoniae (strain R
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: D98066
R.; Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; y, F.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: D98066
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <KUR>
A:Cross-references: GB:AB007317; PIDN:AAL00361.1; PID:g15459222; GSPDB:GN00174
C:Genetics:
A:Gene: spr1557

C;Superfamily: conserved hypothetical protein HI0943

Query Match 83.3%; Score 30; DB 2; Length 177;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 114 KLGRNE 120

RESULT 12

KIECGU

Guanylate kinase (EC 2.7.4.8) - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 01-Mar-2002

C;Accession: S43041; B5166; S24192

R;Gentry, D.; Bangra, C.; Ikehara, K.; Cashel, M.

J. Biol. Chem. 268, 14316-14321, 1993

A;Title: Guanylate kinase of Escherichia coli K-12.

A;Reference number: S43040; MUID:93300828; PMID:8390989

A;Accession: S43041

A;Molecule type: DNA

A;Residues: 1-207 <GEN>

A;Cross-references: EMBL:M84400; NID:q146228; PIDN:AAB89711.1; PID:gl146230

R;Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B5166

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-207 <BLAT>

A;Cross-references: GB:AS000442; GB:U00096; NID:g2367253; PIDN:AAC76672.1; PID:gl790080;

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: gmk

C;Superfamily: guanylate kinase; guanylate kinase homology

C;Keywords: ATP; magnesium; monomer; nucleotide binding; P-loop; phosphotransferase

F;5-187/Domain: guanylate kinase homology <GKI>

F;11-18/Region: nucleotide-binding motif A (P-loop)

F;37-84/Region: GMP binding #status predicted

F;17/Binding site: ATP (Lys) #status predicted

Query Match 83.3%; Score 30; DB 1; Length 207;

Best Local Similarity 85.7%; Pred. No. 41;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 134 RLGRGQ 140

RESULT 13

AI0005

Guanylate kinase (EC 2.7.4.8) [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C;Accession: AI0005

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AI0005

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC8907.1; PID:gl5978155; GSPDB:GN00175

C;Genetics:

A;Gene: gmk

C;Superfamily: guanylate kinase; guanylate kinase homology

C;Keywords: phosphotransferase

Query Match 83.3%; Score 30; DB 2; Length 207;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 134 RLGRGQ 140

RESULT 14

AI0969

Guanylate kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (stra

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AI0969

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serc

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AI0969

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD03253.1; PID:gl6504875; GSPDB:GN00176

C;Genetics:

A;Gene: STY4052

C;Superfamily: guanylate kinase; guanylate kinase homology

Query Match 83.3%; Score 30; DB 2; Length 207;

Best Local Similarity 85.7%; Pred. No. 41;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 134 RLGRGQ 140

RESULT 15

T21262

hypothetical protein F22D6.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21262

R;Wilkinson, J

submitted to the EMBL Data Library, April 1996

A;Reference number: Z19397

A;Accession: T21262

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-445 <WIL>

A;Cross-references: EMBL:Z71262; PIDN:CAA95817.1; GSPDB:GN00019; CESP:F22D6.11

A;Experimental source: clone F22D6

C;Genetics:

A;Gene: CESP:F22D6.11

A;Map position: 1

A;Introns: 65/3; 95/2; 136/3; 180/3; 256/3; 281/3; 403/3

Query Match 83.3%; Score 30; DB 2; Length 445;

Best Local Similarity 71.4%; Pred. No. 87;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 251 RLGRGQ 257

us-09-901-187c-8.rpr

Thu Feb 12 09:07:52 2004

Search completed: February 11, 2004, 17:11:42
Job time : 10.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep:
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep:
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep:
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep:
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep:
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep:
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep:
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep:
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep:
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep:
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep:
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	7	10	US-09-901-187B-8
2	33	91.7	84	12	US-09-864-408A-8020
3	29	80.6	129	15	US-10-106-698-4874
4	29	80.6	696	15	US-10-176-847-46
5	29	80.6	730	15	US-10-217-371-6
6	29	80.6	750	15	US-10-217-371-14
7	29	80.6	751	15	US-10-217-371-4
8	29	80.6	758	15	US-10-217-371-10
9	29	80.6	771	15	US-10-217-371-12
10	29	80.6	779	12	US-10-301-822-151
11	29	80.6	779	15	US-10-217-371-8
12	29	80.6	779	15	US-10-171-311-180
13	29	80.6	790	9	US-09-925-301-1313
14	29	80.6	836	12	US-10-301-822-149
15	29	80.6	836	12	US-10-204-752-33

16	29	80.6	836	12	US-10-295-027-36
17	29	80.6	836	12	US-10-173-989-66
18	29	80.6	836	15	US-10-217-371-2
19	29	80.6	836	15	US-10-171-311-178
20	29	80.6	836	15	US-10-177-293-342
21	29	80.6	2610	12	US-10-369-493-6553
22	29	80.6	2610	12	US-10-369-493-6554
23	28	77.8	54	11	US-09-832-355A-114
24	28	77.8	168	12	US-10-094-749-2002
25	28	77.8	169	12	US-10-114-153-72
26	28	77.8	199	12	US-10-136-728-52
27	28	77.8	199	15	US-10-156-761-8025
28	28	77.8	289	9	US-09-925-301-897
29	28	77.8	313	12	US-10-017-161-752
30	28	77.8	313	12	US-10-041-615-74
31	28	77.8	313	12	US-10-041-615-81
32	28	77.8	346	12	US-10-292-798-656
33	28	77.8	317	10	US-09-886-055-297
34	28	77.8	317	11	US-09-804-291-237
35	28	77.8	469	12	US-10-108-260A-2298
36	28	77.8	474	12	US-10-369-493-2299
37	28	77.8	503	15	US-10-225-060-11
38	28	77.8	506	15	US-10-156-761-8426
39	28	77.8	509	15	US-10-215-224-2
40	28	77.8	509	15	US-10-215-224-3
41	28	77.8	509	15	US-10-225-060-10
42	28	77.8	1038	12	US-10-369-493-1708
43	28	77.8	1134	14	US-10-001-873-50
44	28	77.8	1253	12	US-10-369-493-20694
45	28	77.8	1257	10	US-09-738-626-4750

ALIGNMENTS

RESULT 1
US-09-901-187B-8
; Sequence 8, Application US/09901187B
; Patent No. US200201514641
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Michael S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment of
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-8

Query Match 100.0%; Score 36; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
DB 1 RLGRNQ 7

RESULT 2
US-09-864-408A-8020
; Sequence 8020, Application US/09864408A

```
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8020
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-864-408A-8020

Query Match      91.7%; Score 33; DB 12; Length 84;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
|:|||||
Db 31 RVRGRNQ 37

RESULT 3
US-10-106-698-4874
; Sequence 4874, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4874
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-106-698-4874

Query Match      80.6%; Score 29; DB 15; Length 129;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
|:|||||
Db 4 RVRGRNR 10

RESULT 4
US-10-176-847-46
; Sequence 46, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-176-847-46

Query Match      80.6%; Score 29; DB 15; Length 696;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
|:|||||
Db 33 RIRGRDQ 39

RESULT 5
US-10-217-371-6
; Sequence 6, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-217-371-6

Query Match      80.6%; Score 29; DB 15; Length 730;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
|:|||||
Db 12 RIRGRDQ 18

RESULT 6
US-10-217-371-14
; Sequence 14, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-217-371-14

Query Match      80.6%; Score 29; DB 15; Length 750;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
|:|||||
Db 12 RIRGRDQ 18

RESULT 6
US-10-217-371-14
; Sequence 14, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-217-371-14

Query Match      80.6%; Score 29; DB 15; Length 750;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
|:|:|:|:
Db 12 RIRGRDQ 18

RESULT 7

US-10-217-371-4
; Sequence 4, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ian Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT FILING DATE: 2002-08-13
; PRIOR FILING DATE: 2001-08-13
; CURRENT APPLICATION NUMBER: US 60/312,123
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-4

Query Match 80.6%; Score 29; DB 15; Length 751;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 8

US-10-217-371-10
; Sequence 10, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ian Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT FILING DATE: 2002-08-13
; PRIOR FILING DATE: 2001-08-13
; CURRENT APPLICATION NUMBER: US 60/312,123
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-10

Query Match 80.6%; Score 29; DB 15; Length 751;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 10

US-10-301-822-151
; Sequence 151, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Morahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF COLON CANCER
; FILE REFERENCE: MP01-029P2RNM/10/301,822
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-301-822-151

Query Match 80.6%; Score 29; DB 12; Length 779;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 9

US-10-217-371-12

RESULT 11

US-10-217-371-8
 ; Sequence 8, Application US/10217371
 ; Publication No. US20030073137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Lan Bo
 ; APPLICANT: Dai, Meiru
 ; APPLICANT: Sasaki, Hidesumi
 ; APPLICANT: Auclair, Daniel
 ; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
 ; FILE REFERENCE: 00530-099001
 ; CURRENT APPLICATION NUMBER: US/10/217,371
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US 60/312,123
 ; PRIOR FILING DATE: 2001-08-13
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 779
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-217-371-8

Query Match 80.6%; Score 29; DB 15; Length 779;
 Best Local Similarity 71.4%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 : : : : :
 Db 33 RIRGRDQ 39

RESULT 12
 US-10-171-311-180
 ; Sequence 180, Application US/10171311
 ; Publication No. US20030087270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Chen, Yan
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Monahan, John
 ; APPLICANT: Kamatkar, Shubhang
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Gannavarapu, Manjula
 ; APPLICANT: Hoerth, Sebastian
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
 ; TITLE OF INVENTION: OF CERVICAL CANCER
 ; FILE REFERENCE: MRI-035
 ; CURRENT APPLICATION NUMBER: US/10/171,311
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US 60/298,159
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,155
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/335,936
 ; PRIOR FILING DATE: 2001-11-14
 ; NUMBER OF SEQ ID NOS: 238
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 180
 ; LENGTH: 779
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-171-311-180

Query Match 80.6%; Score 29; DB 15; Length 779;
 Best Local Similarity 71.4%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 : : : : :
 Db 33 RIRGRDQ 39

RESULT 13
 US-09-925-301-1313
 ; Sequence 1313, Application US/09925301
 ; Patent No. US20020052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1313
 ; LENGTH: 790
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-925-301-1313

Query Match 80.6%; Score 29; DB 9; Length 790;
 Best Local Similarity 71.4%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 : : : : :
 Db 44 RIRGRDQ 50

RESULT 14
 US-10-301-822-149
 ; Sequence 149, Application US/10301822
 ; Publication No. US20030148410A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Berger, Allison
 ; APPLICANT: Guillemette, Tracy L.
 ; APPLICANT: Kamatkar, Shubhang
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Thibodeau, Stephen N.
 ; APPLICANT: Burgart, Lawrence J.
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF COLON CANCER
 ; FILE REFERENCE: MPM01-029P2RNM
 ; CURRENT APPLICATION NUMBER: US/10/301,822
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 60/339,971
 ; PRIOR FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: US 60/361,978
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/381,988
 ; PRIOR FILING DATE: 2002-05-20
 ; NUMBER OF SEQ ID NOS: 228
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 149
 ; LENGTH: 836
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-301-822-149

Query Match 80.6%; Score 29; DB 12; Length 836;
 Best Local Similarity 71.4%; Pred. No. 1.3e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 : : : : :
 Db 33 RIRGRDQ 39

Thu Feb 12 09:07:51 2004

RESULT 15
US-10-204-752-33
; Sequence 33, Application US/10204752
; Publication No. US20030152956A1
; GENERAL INFORMATION:
; APPLICANT: OHTANI, No. US20030152956A1iko
; APPLICANT: MATSUI, Keiko
; APPLICANT: YOSHIDA, Nei
; APPLICANT: SUGITA, Yuji
; APPLICANT: IZUHARA, Kenji
; TITLE OF INVENTION: METHOD OF TESTING FOR ALLERGIC DISEASE
; FILE REFERENCE: SHZ-009US
; CURRENT APPLICATION NUMBER: US/10/204,752
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: JP 2000-396166
; PRIOR FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 836
; TYPE: PRI
; ORGANISM: Homo sapiens
US-10-204-752-33

Query Match 80.6%; Score 29; DB 12; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RLGRNQ 7
|:|:|:
Db 33 RIQRDQ 39

Search completed: February 11, 2004, 17:54:11
Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-8

Perfect score: 36

Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	88.9	494	US-09-252-991A-17295	Sequence 17295, A
2	30	83.3	195	US-09-252-991A-16682	Sequence 16682, A
3	30	83.3	281	US-09-252-991A-33048	Sequence 33048, A
4	30	83.3	357	US-09-252-991A-31718	Sequence 31718, A
5	30	83.3	435	US-09-252-991A-30142	Sequence 30142, A
6	30	83.3	528	US-09-252-991A-28291	Sequence 28291, A
7	29	80.6	779	US-08-436-627-4	Sequence 4, Appl
8	29	80.6	779	US-08-436-627-24	Sequence 24, Appl
9	29	80.6	811	US-08-436-627-2	Sequence 2, Appl
10	29	80.6	811	US-08-436-627-22	Sequence 22, Appl
11	29	80.6	836	US-08-436-627-6	Sequence 6, Appl
12	29	80.6	837	US-08-436-627-23	Sequence 23, Appl
13	28	77.8	260	US-08-879-561-8	Sequence 8, Appl
14	28	77.8	262	US-09-857-498A-6	Sequence 6, Appl
15	28	77.8	310	US-09-107-532A-4212	Sequence 4212, Ap
16	28	77.8	442	US-09-252-991A-31848	Sequence 31848, A
17	28	77.8	503	US-08-740-223A-11	Sequence 11, Appl
18	28	77.8	503	US-09-709-188-11	Sequence 11, Appl
19	28	77.8	509	US-08-665-926-8	Sequence 8, Appl
20	28	77.8	509	US-08-740-223A-10	Sequence 10, Appl
21	28	77.8	509	US-09-202-491-2	Sequence 2, Appl
22	28	77.8	509	US-09-202-491-3	Sequence 3, Appl
23	28	77.8	509	US-09-709-188-10	Sequence 10, Appl
24	28	77.8	736	US-09-102-528-29	Sequence 29, Appl
25	28	77.8	1038	US-09-541-762-4	Sequence 4, Appl
26	28	77.8	1038	US-09-723-820-4	Sequence 4, Appl
27	28	77.8	1184	US-09-541-762-2	Sequence 2, Appl

Sequence 2, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 109, App
Sequence 110, App
Sequence 109, App
Sequence 110, App
Sequence 109, App
Sequence 110, App
Sequence 99, Appl
Sequence 99, Appl
Sequence 99, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 8, Appl

28 77.8 1184 4 US-09-723-820-2
29 77.8 1257 2 US-08-750-152A-2
30 43 1 US-08-098-005-1
31 73 1 US-07-609-716-109
32 73 1 US-07-609-716-110
33 73 1 US-08-475-411A-109
34 73 3 US-08-475-411A-110
35 73 3 US-08-478-029A-109
36 73 3 US-08-478-029A-110
37 73 1 US-07-609-716-99
38 73 97 3 US-08-475-411A-99
39 73 97 3 US-08-478-029A-99
40 99 2 US-08-292-968-25
41 99 2 US-08-467-974-25
42 99 2 US-08-467-536-25
43 99 3 US-08-467-976-25
44 99 3 US-09-082-514-25
45 100 2 US-08-949-376-8

ALIGNMENTS

RESULT 1
US-09-252-991A-17295
; Sequence 17295, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17295
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17295

Query Match 88.9%; Score 32; DB 4; Length 494;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
DB 243 RLGRNR 249

RESULT 2
US-09-252-991A-16682
; Sequence 16682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16682
; LENGTH: 195
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16682

Query Match 83.3%; Score 30; DB 4; Length 195;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 21 RLGRGQ 27

RESULT 3

US-09-252-991A-33048
; Sequence 33048, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 33048

; LENGTH: 291

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-33048

Query Match 83.3%; Score 30; DB 4; Length 281;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 59 RLGRGQ 65

RESULT 4

US-09-252-991A-31718
; Sequence 31718, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31718

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31718

Query Match 83.3%; Score 30; DB 4; Length 357;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 48 RLGRGQ 54

RESULT 5

US-09-252-991A-30142
; Sequence 30142, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30142

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30142

Query Match 83.3%; Score 30; DB 4; Length 435;
Best Local Similarity 85.7%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 254 RLGRGQ 260

RESULT 6

US-09-252-991A-28291
; Sequence 28291, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28291

; LENGTH: 528

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28291

Query Match 83.3%; Score 30; DB 4; Length 528;
Best Local Similarity 85.7%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||
Db 170 RLGRGQ 176

RESULT 7

US-08-426-627-4
; Sequence 4, Application US/08426627
; Patent No. 5756654

; GENERAL INFORMATION:

; APPLICANT: Amanu, Egon

; APPLICANT: Otawara-Hamamoto, Yoko

; APPLICANT: Kikuno, Reiko

; APPLICANT: Takeshita, Sunao

; APPLICANT: Tezuka, Kenichi

;; TITLE OF INVENTION: No. 5756664e1 Protein with Bone Formation
;; TITLE OF INVENTION: Ability and Process for Its Production.
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;; ADDRESSEE: Dunner
;; STREET: 1300 I Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3315
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; FILING DATE: US/08/426,627
;; PRIORITY INFORMATION:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/036,841
;; FILING DATE: 25-MAR-1993
;; APPLICATION NUMBER: JP 4-71501
;; FILING DATE: 27-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hammond, Alan W.
;; REGISTRATION NUMBER: 35,178
;; REFERENCE/DOCKET NUMBER: 02481-1285-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4400
;; TELEFAX: 202-408-4400
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 779 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-426-627-4

Query Match 80.6%; Score 29; DB 1; Length 779;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7
|:|||||
Db 33 RIRGRDQ 39

RESULT 8
US-08-426-627-24
; Sequence 24, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664e1 Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/426,627
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/426,627
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/036,841
;; FILING DATE: 25-MAR-1993
;; APPLICATION NUMBER: JP 4-71501
;; FILING DATE: 27-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hammond, Alan W.
;; REGISTRATION NUMBER: 35,178
;; REFERENCE/DOCKET NUMBER: 02481-1285-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4000
;; TELEFAX: 202-408-4400
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 779 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-426-627-24

Query Match 80.6%; Score 29; DB 1; Length 779;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7
|:|||||
Db 33 RIRGRDQ 39

RESULT 9
US-08-426-627-2
; Sequence 2, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664e1 Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/426,627
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178

REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-426-627-2

Query Match 80.6%; Score 29; DB 1; Length 811;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7
|:|:|:|:
Db 35 RIRGRDQ 41

RESULT 10
US-08-426-627-22
Sequence 22, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-426-627-22

Query Match 80.6%; Score 29; DB 1; Length 811;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RLGRNQ 7
|:|:|:|:
Db 35 RIRGRDQ 41

RESULT 11
US-08-426-627-6
Sequence 6, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 836 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-426-627-6

Query Match 80.6%; Score 29; DB 1; Length 836;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 12
US-08-426-627-23
Sequence 23, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Otawara-Hamamoto, Yoko


```

/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneka
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 4212:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 310 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (B) LOCATION 1...310
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4212:
US-09-107-532A-4212

```

```

Query Match          77.8%; Score 28; DB 4; Length 310;
Best Local Similarity 71.4%; Pred No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 RLGRNQ 7
Db      277 RLRGNE 283

```

```

Search completed: February 11, 2004, 17:13:37
Job time : 10.4167 secs

```

XX
PT Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper
 XX
 PS Claim 40; Page 37; 52pp; English.
 CC The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 Db 1 RLGRNQ 7

RESULT 2
 ABB63409
 ID ABB63409 standard; Protein; 533 AA;

AC ABB63409;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17019.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL07512.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 17019; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16175-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 533 AA;

Query Match 100.0%; Score 36; DB 22; Length 533;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 Db 503 RLGRNQ 509

RESULT 3

ABP35037

ID ABP35037 standard; Protein; 84 AA.

XX ABP35037;

DT 09-JUL-2002 (first entry)

XX Human ORF4010 protein, SEQ ID NO:8020.

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytoagulant; neurotropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiact; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.

PN WO200190366-A2.

PD 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US17076.

XX 24-MAY-2000; 2000US-206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shimkets RA;

XX WPI; 2002-106200/14.

DR N-PSDB; ABL79063.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -

PS Claim 10; Page 2251; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABL75054-
 CC ABL79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide

CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, anti-inflammatory activity, tumour inhibition activity,
CC and antifibrotic activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC other proinflammatory disorders such as psoriasis and benign tumours,
CC other proliferative disorders such as epilepsy and Alzheimer's disease,
CC neurological disorders, immune system disorders, disorders related to
CC cardiovascular diseases, disorders of tissue growth and regeneration,
CC organ transplantation, disorders of hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX SQ Sequence 84 AA;
Query Match 91.7%; Score 33; DB 23; Length 84;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|:|||||
DB 31 RVGRNQ 37

RESULT 4
ABG17799
ID ABG17799 standard; Protein; 458 AA.

XX AC ABG17799;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #17790.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX XX WPI; 2001-639362/73.

DR N-PSDB; AA981986.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX PS Claim 20; SEQ ID No 48158; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II) (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG0377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 458 AA;

Query Match 88.9%; Score 32; DB 22; Length 458;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7

DB 253 QLGRNQ 259

RESULT 5

ABE49607

ID ABE49607 standard; Protein; 75 AA.

XX AC ABE49607;

XX DT 05-FEB-2002 (first entry)

XX DE Listeria monocytogenes protein #2311.

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.

XX OS Listeria monocytogenes.

XX PN WO200177335-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-FR01118.

XX PR 11-APR-2000; 2000FR-0004629.

XX PA (INSP) INST PASTEUR.

XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goshel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durand L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;

XX DR WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -

XX

PS Claim 6; SEQ ID No 2312; 192pp; French.

CC The present invention relates to the genome sequence of *Listeria*

CC monocytes RGD-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in *L.*

CC monocytes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present invention is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of *L. monocytogenes* and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate *L. monocytogenes*-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by *L.*

CC monocytes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 75 AA;

Query Match 86.1%; Score 31; DB 23; Length 75;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6

Db 70 RLGRN 75

RESULT 6

AAU61300

ID AAU61300 standard; Protein: 78 AA.

AC AAU61300;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22196.

XX

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertostis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

XX

OS Propionibacterium acnes.

XX

PN WO200181581-A2.

XX

PD 01-NOV-2001.

XX

XX 20-APR-2001; 2001WO-US12865.

XX

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX

XX (CORI-) CORIXA CORP.

XX

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

PI

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59616.

DR

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX

PS Example 1; SEQ ID No 22495; 1069pp; English.

XX

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, imititis,

CC pustulosis, hypertostis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 78 AA;

Query Match 86.1%; Score 31; DB 22; Length 78;

Best Local Similarity 85.7%; Pred. No. 34;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7

Db 47 RLGRGDQ 53

RESULT 7

ABG20695

ID ABG20695 standard; Protein: 87 AA.

XX

AC ABG20695;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #20686.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW Homo sapiens.

OS

XX

XX WO200175067-A2.

PN

XX

PD 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US08631.

XX

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

XX (HYSE-) HYSEQ INC.

PA

XX

XX Drmanac RT, Liu C, Tang YT;

PI

XX WPI; 2001-639362/73.

DR N-PSDB; AAS84882.

DR

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

XX Claim 20; SEQ ID No 51054; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 87 AA;

Query Match 86.1%; Score 31; DB 22; Length 87;

Best Local Similarity 85.7%; Pred. No. 38;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7

|||||

32 RLGRSQ 38

RESULT 8

AA18501
ID AAG18501 standard; Protein; 131 AA.

AC AAG18501;

XX

XX

DT 17-OCT-2000 (first entry)

XX

DE Zea mays protein fragment SEQ ID NO: 19935.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.

XX

OS Zea mays subsp. mays.

XX

XX

FN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0121180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125786.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 18-APR-1999; 99US-0128714.

PR

PR 08-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 28-APR-1999; 99US-0130891.

PR

PR 30-APR-1999; 99US-0131449.

PR

PR 10-APR-1999; 99US-0132048.

PR

PR 04-MAY-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 06-MAY-1999; 99US-0132487.

PR

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 14-MAY-1999; 99US-0134768.

PR 18-MAY-1999; 99US-0134941.

PR 19-MAY-1999; 99US-0135124.

PR 20-MAY-1999; 99US-0135353.

PR 21-MAY-1999; 99US-0135629.

PR 24-MAY-1999; 99US-0136021.

PR 25-MAY-1999; 99US-0136392.

PR 27-MAY-1999; 99US-0136782.

PR 28-MAY-1999; 99US-0137222.

PR 01-JUN-1999; 99US-0137528.

PR 03-JUN-1999; 99US-0137502.

PR 04-JUN-1999; 99US-0137724.

PR 07-JUN-1999; 99US-0138094.

PR 08-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 10-JUN-1999; 99US-0139119.

PR 14-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 16-JUN-1999; 99US-0139492.

PR 17-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139465.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 18-JUN-1999; 99US-0139817.

PR 21-JUN-1999; 99US-0139899.

PR 22-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 05-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0143220.

PR 12-JUL-1999; 99US-0143277.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 18-JUL-1999; 99US-0144086.

PR 18-JUL-1999; 99US-0144325.

PR 18-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

```
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 131;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
Db 8 RLGRN 13

RESULT 9
AAG25683
ID AAG25683 standard; Protein; 139 AA.
XX AC AAG25683;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 29845.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
```

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139455.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140355.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0134941.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 139;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
Db 8 RLRGRN 13
|||||

RESULT 10
AAG18500
ID AAG18500 standard; Protein; 145 AA.
XX
AC AAG18500;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 19934.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135553.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139829.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

Query Match 86.1%; Score 31; DB 21; Length 145;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RLRGRN 6
Db 22 RLRGRN 27

RESULT 11
AAG25682
ID AAG25682 standard; Protein; 153 AA.
XX AC AAG25682;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 29844.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 16-APR-1999; 99US-0128714.
XX PR 19-APR-1999; 99US-0129845.
XX PR 21-APR-1999; 99US-0130077.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 18-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.

RESULT 12
AAG16627
ID AAG16627 standard; Protein; 181 AA.
XX AC AAG16627;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17345.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX DD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136382.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.


```

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 161;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
Db 8 RLRGRN 13

RESULT 13
AAG16626
ID AAG16626 standard; Protein; 195 AA.
XX
AC AAG16626;
```

```

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17344.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137242.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
```

```
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159333.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. NO. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
Db 22 RLGRN 27

RESULT 14
ABB50063
ID ABB50063 standard; Protein; 280 AA.
XX
AC ABB50063;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #2767.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
vitamin B12; bacterial infection; disease.
```

XX OS Listeria monocytogenes.
XX OS WC200177335-A2.
XX PN
XX PD 19-OCT-2001.
XX XX
XX PF 11-APR-2001; 2001WO-FR01118.
XX XX
XX PR 11-APR-2000; 2000FR-0004629.
XX XX
XX PA (INSP) INST PASTEUR.
XX XX
XX PI Buchrieser C, Frangeul L, Couve E, Ruenick C, Fsihi H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Krefit J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX XX
XX WPI; 2002-010914/01.
XX XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and
XX PT related polypeptides -
XX XX
XX Claim 6; SEQ ID No 2768; 192pp; French.
XX XX
XX The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded from the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX Sequence 280 AA;
XX SQ
Query Match 86.1%; Score 31; DB 23; Length 280;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RLGRNQ 7
DB 192 RLHGRNQ 198
RESULT 15
AAU46599
ID AAU46599 standard; Protein; 337 AA.
XX XX
XX AC AAU46599;
XX XX
XX DT 27-FEB-2002 (first entry)
XX XX
XX DE Propionibacterium acnes immunogenic protein #7495.
XX XX
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

dermatological; osteopathic; neuroprotectant.
Propionibacterium acnes.
WO200181581-A2.
01-NOV-2001.
20-APR-2001; 2001WO-US12865.
21-APR-2000; 2000US-199047P.
02-JUN-2000; 2000US-20841P.
07-JUL-2000; 2000US-216747P.
(CORI-) CORIXA CORP.
Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
L'maisonneuve J, Zhang Y, Jen S, Carter D;
WPI; 2001-616774/71.
N-PSDB; AAS59534.
Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris -
Example 1; SEQ ID No 7794; 1069pp; English.
Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
polypeptides. The proteins and their associated DNA sequences are used in
the treatment, prevention and diagnosis of medical conditions caused by
P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
P. acnes is also involved in infections of bone, joints and the central
nervous system, however it is particularly involved in the inflammatory
lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting a
sample with a binding agent that binds to the proteins of the invention
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibodies
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be used as
diagnostic agents for determining P. acnes presence, for example, by
enzyme linked immunosorbent assay (ELISA).
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX Sequence 337 AA;
XX SQ
Query Match 86.1%; Score 31; DB 22; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LRGRNQ 7
DB 136 LRGRNQ 141
Search completed: February 11, 2004, 17:02:52
Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL 23: *
2: sp_archea: *
3: sp_bacteria: *
4: sp_fungi: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	196	5 Q9U3C2	Q9U3C2 caenorhabdi
2	42	79.2	272	13 Q8JH73	Q8JH73 xenopus lae
3	42	79.2	309	10 Q49746	Q49746 arabidopsis
4	42	79.2	310	10 Q9M2Y9	Q9M2Y9 arabidopsis
5	42	79.2	487	16 Q9CKV7	Q9CKV7 pasteurella
6	42	79.2	2197	5 Q9W016	Q9W016 drosophila
7	40	75.5	732	13 Q9GU61	Q9GU61 myxine glut
8	40	75.5	749	11 Q8BQ07	Q8BQ07 mus musculu
9	40	75.5	1826	4 Q8IZX4	Q8IZX4 homo sapien
10	40	75.5	1865	11 Q60544	Q60544 mesocricetu
11	39	73.6	179	11 Q8CUD2	Q8CUD2 mus musculu
12	39	73.6	242	10 Q9ARQ2	Q9ARQ2 cryza sativ
13	39	73.6	375	6 Q97505	Q97505 sus scrofa
14	39	73.6	563	2 Q8RS37	Q8RS37 pseudomonas
15	39	73.6	750	10 Q81880	Q81880 arabidopsis
16	39	73.6	826	10 Q940X9	Q940X9 arabidopsis

17 38.5 72.6 750 3 O42657
18 38 71.7 193 10 Q9LWM8
19 38 71.7 284 16 O06252
20 38 71.7 293 3 Q12393
21 38 71.7 312 16 Q8VC08
22 38 71.7 316 16 Q9KS69
23 38 71.7 370 10 Q9LGV4
24 38 71.7 381 8 Q8HHQ4
25 38 71.7 412 10 Q8RYL7
26 38 71.7 545 2 Q9F3V6
27 38 71.7 545 2 Q9F3V6
28 38 71.7 708 16 Q8CZY0
29 38 71.7 708 16 Q8CZY0
30 38 71.7 753 16 Q8ZHU0
31 37.5 70.8 766 16 O06817
32 37.5 70.8 452 4 Q96AB7
33 37.5 70.8 484 4 Q9BTV6
34 37 69.8 100 12 Q9YRJ4
35 37 69.8 101 12 Q9YRJ7
36 37 69.8 105 12 Q9YRJ6
37 37 69.8 111 12 Q9YRJ8
38 37 69.8 112 12 Q9YRJ7
39 37 69.8 119 9 Q8LTS5
40 37 69.8 138 12 Q9YRI5
41 37 69.8 139 12 Q9YRJ2
42 37 69.8 142 10 Q94J29
43 37 69.8 221 3 Q9F8K9
44 37 69.8 299 10 Q8RWG4
45 37 69.8 313 10 Q9MOM0

ALIGNMENTS

RESULT 1

Q9U3C2 PRELIMINARY; PRT; 196 AA.
ID Q9U3C2
AC Q9U3C2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE K03H1.9 protein.
GN K03H1.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith M.J., Ainscough R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z29550; CAA82661.2; -.
DR WormPep; K03H1.9; CE23851.
SQ SEQUENCE 196 AA; 24273 MW; 958E7B149C150D4C CRC64;

Query Match 79.2%; Score 42; DB 5; Length 196;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFFHHR 7
|||
DB 78 PFFHHR 83

RESULT 2

Q8JH73

```

ID Q9JH73 PRELIMINARY; PRT; 272 AA.
AC Q9JH73
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integral membrane nucleoporin gp210 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22100007; PubMed=12093788;
RA Drummond S.P., Wilson K.L.;
RT "Interference with the cytoplasmic tail of gp210 disrupts 'close
RL apposition' of nuclear membranes and blocks nuclear pore dilation.";
RL J. Cell Biol. 158:53-62(2002).
DR EMBL; AF533550; AAM94631.1; -.
KW Porin
FT NON-TER
SQ SEQUENCE 272 AA; 29952 MW; EE9023786FF9988 CRC64;

Query Match 79.2%; Score 42; DB 13; Length 272;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHH 7
DB 225 PFHHH 230

RESULT 3
C49746 PRELIMINARY; PRT; 309 AA.
ID C49746
AC C49746
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE R2R3-MYB transcription factor.
GN MYB84.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Romero I., Fuentes A., Benito M.J., Malpica J., Leyva A., Paz-Ares J.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL; Y14209; CAA74605.1; -.
DR EMBL; Z95807; CAB09239.1; -.
DR HSSP; P06876; 1MBX.
DR TRANSFAC; T02597; -.
DR InterPro; IPR001005; MYB DNA binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS00090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 309 AA; 35545 MW; 8E673F408803007C CRC64;

Query Match 79.2%; Score 42; DB 10; Length 309;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
DB 163 WPLHHH 168

RESULT 4
Q9M2Y9 PRELIMINARY; PRT; 310 AA.
ID Q9M2Y9
AC Q9M2Y9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AtMYB84.
GN T16K5.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA Mewes H.W., Lemcke K., Mayer K.F.X., Quettier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL; AL132965; CAB6907.1; -.
DR HSSP; P06876; 1MBX.
DR InterPro; IPR001005; MYB DNA binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS00090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 310 AA; 35578 MW; 647FB85D226CB6C3 CRC64;

Query Match 79.2%; Score 42; DB 10; Length 310;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
DB 163 WPLHHH 168

RESULT 5
Q9CKV7 PRELIMINARY; PRT; 487 AA.
ID Q9CKV7
AC Q9CKV7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TrkH.
GN TRKH OR PM1498.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OC NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006186; AA03582.1; -.
DR InterPro; IPR003445; Cat_transpt.
DR InterPro; IPR004772; K_transpttrk.
DR Pfam; PF02386; TrkH; 1.
DR TIGRFAMs; TIGR00933; 2a38; 1.
KW Complete proteome.
SQ SEQUENCE 487 AA; 54000 MW; 15049745C5872BDE CRC64;

```

```

Query Match          79.2%; Score 42; DB 16; Length 487;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 WPFHHR 7
DB 55 WPCHHK 61

RESULT 6
Q9W016 PRELIMINARY; PRT; 2197 AA.
ID OSW016;
AC OSW016;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG32306 protein.
GN CG32306 OR CG13790 OR CG13797 OR CG13798 OR CG13799 OR CG16763.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.E., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,

```

```

RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003474; AAF47644.2; -
DR FlyBase; FBGN0052306; CG32306;
SQ SEQUENCE 2197 AA; 237021 MW; 069D7AC05B10AAEB CRC64;

Query Match          79.2%; Score 42; DB 5; Length 2197;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 6
DB 336 WPLHHH 341

RESULT 7
Q9GU61 PRELIMINARY; PRT; 732 AA.
ID Q9GU61;
AC Q9GU61;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RING3 (Fragment).
GN RING3.
OS Myxine glutinosa (Atlantic hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxiniinae; Myxine.
OX NCBI_TaxID=7769;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Peripheral blood leukocytes;
RA White G.P., Cunningham C.;
RT "Characterization of RING3-like protein from Atlantic hagfish (Myxine glutinosa)".
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF191032; AAG1179.1; -
DR HSSP; Q92831; 1B91.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00634; BROMODOMAIN_2; 2.
FT NON_TER 732
SQ SEQUENCE 732 AA; 81669 MW; 2038D61BD20F6B50 CRC64;

```

Query Match 75.5%; Score 40; DB 13; Length 732;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||
DB 53 WPFHH 57

RESULT 8

Q8BQ07 PRELIMINARY; PRT; 749 AA.
AC Q8BQ07;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE TAFI1250 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK046668; BAC32828.1; -.
FT NON TER 1
SQ SEQUENCE 749 AA; 85475 MW; 2PAF7354AEC42710 CRC64;

Query Match 75.5%; Score 40; DB 11; Length 749;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||
DB 405 WPFHH 409

RESULT 9

Q81ZX4 PRELIMINARY; PRT; 1826 AA.
AC Q81ZX4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE TBP-associated factor RNA polymerase 1-like.
GN TAF1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22206412; PubMed=12217962;
RA Wang P.J., Page D.C.;
RT "Functional substitution for TAF(II)250 by a retroposed homolog that
RT is expressed in human spermatogenesis.";
RL Hum. Mol. Genet. 11:2341-2346(2002).
DR EMBL; AF390562; AAN40840.1; -.
SQ SEQUENCE 1826 AA; 207300 MW; 35D780E749AC9B17 CRC64;

Query Match 75.5%; Score 40; DB 4; Length 1826;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||

DB 1545 WPFHH 1549

RESULT 10

Q60544 PRELIMINARY; PRT; 1865 AA.
AC Q60544;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CCG1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94215915; PubMed=8163200;
RA Hayashida T., Sekiguchi T., Noguchi E., Sunamoto H., Ohba T.,
RA Nishimoto T.,
RT "The CCG1/TAFI1250 gene is mutated in thermosensitive G1 mutants of
RT the BHK21 cell line derived from goldenhamster.";
RL Gene 141:267-270(1994).
DR EMBL; D26114; BAA05110.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 2.
DR PROSITE; PS00014; BROMODOMAIN 2; 2.
SQ SEQUENCE 1865 AA; 211866 MW; A81614946C0C0F24 CRC64;

Query Match 75.5%; Score 40; DB 11; Length 1865;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||
DB 1521 WPFHH 1525

RESULT 11

Q9CUD2 PRELIMINARY; PRT; 179 AA.
AC Q9CUD2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 4930520K1ORik protein (Fragment).
GN 4930520K1ORIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Montbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690 (2001).
 RL EMBL; AK016752; BAB30410.1; -.
 DR MGD; MGI:1914837; 4930520K10Rik.
 FT NON_TER 1
 SQ SEQUENCE 179 AA; 18465 MW; 583C70B92085CE9 CRC64;

Query Match 73.6%; Score 39; DB 11; Length 179;
 Best Local Similarity 71.4%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHHR 7
 DB 6 WKFHHR 12

RESULT 12

ID Q9ARQ2 PRELIMINARY; PRT; 242 AA.
 AC Q9ARQ2; 2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE OSJNBAC010K01.13 protein.
 GN OSJNBAC010K01.13.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 CX NCBI_TaxID=4530;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone:OSJNBAC010K01.1";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003210; BAB40099.1; -.
 DR Gramene; Q9ARQ2; -.
 DR InterPro; IPR007087; Znf C2H2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 SQ SEQUENCE 242 AA; 27507 MW; EAE4EA06AA06D785 CRC64;

Query Match 73.6%; Score 39; DB 10; Length 242;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 6
 DB 51 WPGHHH 56

RESULT 13

ID Q97505 PRELIMINARY; PRT; 375 AA.
 AC Q97505;
 DT 01-MAY-1999 (TREMELrel. 10, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Neurotrophin Y receptor type 4 (Neurotrophin Y-family receptor
 DE Y4).
 GN NPY Y4.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LWD; TISSUE=Kidney;

RA Ito Y., Minezawa M.;
 RT "Sus scrofa NPY Y4 gene for neuropeptide Y receptor type 4, complete
 RT cds [genomic].";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BREED NORWEGIAN LANDRACE;
 RX MEDLINE=20187975; PubMed=10720571;
 RA Wraith A., Turnstin A., Chardon P., Harbitz I., Chowdhary B.P.,
 RA Andersson L., Lundin L.-G., Larhammar D.;
 RT "Evolution of the neuropeptide Y receptor family: gene and chromosome
 RT duplications deduced from the cloning and mapping of the five receptor
 RT subtype genes in the pig";
 RL Genome Res. 10:302-310 (2000).
 DR EMBL; AB021678; BAA36218.1; -.
 DR EMBL; AF227955; AAF62507.1; -.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW Neuropeptide; Receptor.
 SQ SEQUENCE 375 AA; 42346 MW; 5992F075904DB1A0 CRC64;

Query Match 73.6%; Score 39; DB 6; Length 375;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
 DB 205 WPLEHHR 211

RESULT 14

Q8RS37 PRELIMINARY; PRT; 563 AA.
 AC Q8RS37;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE WD-repeat protein.
 GN TOXC.
 OS Pseudomonas glumae.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 CX NCBI_TaxID=337;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF327195;
 RA Suzuki F.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF327195;
 RA Suzuki F., Sawada H., Matsuda I.;
 RT "Molecular Characterization of Toxoflavin Biosynthesis-related Gene in
 RT Pseudomonas (Burkholderia) glumae";
 RL Ann. Phytopathol. Soc. Jpn. 64:276-281 (1998).
 DR EMBL; AB040403; BAB88914.1; -.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 6.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00578; WD_REPEATS_1; 1.
 DR PROSITE; PS0082; WD_REPEATS_2; 4.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 563 AA; 61957 MW; 6636126F0590931D CRC64;

Query Match 73.6%; Score 39; DB 2; Length 563;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7

Db 501 WPLHR 507

RESULT 15
 O81880 PRELIMINARY; PRT; 750 AA.
 AC O81880;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical 83.9 kDa protein.
 GN T16L1.120 OR A1933630.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Deutschenbaur S., Piravandi E., Hoheisel J., Jesse T.,
 RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Deutschenbaur S., Piravandi E., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031394; CAA20576.1;
 DR EMBL; AL161583; CAB80080.1;
 KW Hypothetical protein.
 SQ SEQUENCE 750 AA; 83907 MW; 3F298ECB1FD505A6 CRC64;

Query Match 73.6%; Score 39; DB 10; Length 750;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 PFHHR 7
 Db 21 PYHHR 26

Search completed: February 11, 2004, 17:09:41
 Job time : 33.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-9

Perfect score: 53
Sequence: 1 WFFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	79.2	497	1 TRKH_HAEN	P44843 haemophilus
2	40	75.5	1872	1 T2D1_HUMAN	P21675 homo sapien
3	39	73.6	316	1 COAA_BACHD	Q9K8X7 bacillus ha
4	38	71.7	744	1 BISC_RHOSH	P54934 rhodobacter
5	38	71.7	925	1 PIP1_YEAST	P40020 saccharomyc
6	37	69.8	221	1 HEX1_EMENI	Q9P8K9 smericella
7	37	69.8	236	1 MRAY_ARATH	O49730 arabidopsis
8	37	69.8	327	1 COL7_ARATH	Q9SYM2 arabidopsis
9	37	69.8	336	1 MRAY_STRP3	QAK6C7 streptococc
10	37	69.8	336	1 MRAY_STRP8	QGNZY2 streptococc
11	37	69.8	336	1 MRAY_STRP9	Q9YXK2 streptococc
12	37	69.8	359	1 MRAY_AQUAE	O66465 aquifex aeo
13	37	69.8	359	1 MRAY_MYCLE	O69555 mycobacteri
14	37	69.8	359	1 MRAY_MYCTU	O06221 mycobacteri
15	37	69.8	375	1 NY4R_HUMAN	P50391 homo sapien
16	37	69.8	395	1 DITE_BACSU	P39580 bacillus su
17	37	69.8	611	1 THIC_RHIME	Q22T19 rhizobium m
18	37	69.8	619	1 THIC_RHILLO	Q98A23 rhizobium l
19	37	69.8	1849	1 T172_HUMAN	O14981 homo sapien
20	36	67.9	404	1 ZFS1_SCHPO	P47979 schizosacch
21	36	67.9	526	1 CNA2_YEAST	P06776 saccharomyc
22	35	66.0	254	1 SZA2_HUMAN	P31213 homo sapien
23	35	66.0	254	1 SZA2_MACFA	Q28892 macaca fasc
24	35	66.0	254	1 SZA2_FIG	O18765 sus scrofa
25	35	66.0	319	1 TCBI_FABIT	P06333 cryocologus
26	35	66.0	356	1 POU1_XENLA	P31363 xenopus lae
27	35	66.0	386	1 HXAD_MOUSE	Q62424 mus musculu
28	35	66.0	388	1 HXAD_HUMAN	P31271 homo sapien
29	35	66.0	447	1 HEX_ADE04	P36850 human adeno
30	35	66.0	457	1 THIC_ANASP	Q8YV69 anabaena sp
31	35	66.0	459	1 THIC_SYNZ3	Q55894 synechocyst
32	35	66.0	465	1 HEX_ADE06	Q04966 human adeno
33	35	66.0	467	1 HEX_ADE01	Q04965 human adeno

34	35	66.0	468	1 HEX_ADE31	P36855 human adeno
35	35	66.0	501	1 FXC2_HUMAN	Q99958 homo sapien
36	35	66.0	512	1 HEX_ADE09	P36853 human adeno
37	35	66.0	517	1 HEX_ADE08	P36852 human adeno
38	35	66.0	905	1 HEX_ADECC	Q65955 canine aden
39	35	66.0	905	1 HEX_ADECR	O39619 canine aden
40	35	66.0	909	1 HEX_ADEW1	P48308 mouse adeno
41	35	66.0	911	1 HEX_ADEB3	P03278 bovine adeno
42	35	66.0	919	1 HEX_ADE12	P19900 human adeno
43	35	66.0	923	1 HEX_ADE40	P11819 human adeno
44	35	66.0	925	1 HEX_ADE41	P11820 human adeno
45	35	66.0	937	1 HEX_ADE07	P36851 human adeno

ALIGNMENTS

RESULT 1
TRKH_HAEN
ID TRKH_HAEN STANDARD; PRT; 487 AA.
AC P44843;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trk system potassium uptake protein trkh.
GN TRKH OR H10723.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RD / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD.";
RL Science 269:496-512(1995).
RN [2]
RP REVISIONS.
RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.,
RA Peterson J., Hickey E., Dodson R., Gwinn M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: LOW-AFFINITY POTASSIUM TRANSPORT SYSTEM. INTERACTS WITH
CC TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA AND REQUIRES TRKE FOR
CC TRANSPORT ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U32755; AAC22381.1; --
CC TIGR; H10723; --
CC InterPro; IPR003445; Cat.transpr.
CC InterPro; IPR004772; K.transprtrk.
CC Pfam; PF02386; Trkh; 1.
CC TIGRfams; TIGR00933; 2a38; 1.
CC Transmembrane; Inner membrane; Transport; Potassium transport;
CC Complete proteome.
KW

```

FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 401 421 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 461 481 POTENTIAL.
SQ SEQUENCE 487 AA; 53458 MW; 91A8P38C37P4540A CRC64;

Query Match 79.2%; Score 42; DB 1; Length 487;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
Db 55 WPCHHK 61

RESULT 2
T2D1_HUMAN STANDARD; PRT; 1872 AA.
AC P21675;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription initiation factor TFIIID 250 kDa subunit (TAFII-250)
DE (TAFII250) (TBP-associated factor 250 kDa) (P250) (Cell cycle gene 1
DE protein)
DE TAFI OR TAF2A OR CCG1 OR B42R.
GN Homo sapiens (Human).
OS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Laryngeal carcinoma;
RX MEDLINE=91246200; PubMed=2038334;
RA Sekiguchi T., Nohiro Y., Nakamura Y., Hisamoto N., Nishimoto T.;
RT "The human CCG1 gene, essential for progression of the G1 phase,
RT encodes a 210-kilodalton nuclear DNA-binding protein.";
RL Mol. Cell. Biol. 11:3317-3325(1991).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RX MEDLINE=89005056; PubMed=3169001;
RA Sekiguchi T., Miyata T., Nishimoto T.;
RT "Molecular cloning of the cDNA of human X chromosomal gene (CCG1)
RT which complements the temperature-sensitive G1 mutants, tsN462 and
RT ts13, of the BHK cell line.";
RL EMBO J. 7:1683-1687(1988).
RN [3]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=93196704; PubMed=7680771;
RA Ruppert S., Wang E.H., Tian R.;
RT "Cloning and expression of human TAFII250: a TBP-associated factor
RT implicated in cell-cycle regulation.";
RL Nature 362:175-179(1993).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=93196705; PubMed=8450888;
RA Hisatake K., Hasegawa S., Takada R., Nakatani Y., Horikoshi M.,
RA Roeder R.G.;
RT "The p250 subunit of native TATA box-binding factor TFIID is the
RT cell-cycle regulatory protein CCG1.";
RL Nature 362:179-181(1993).
RN
CC -1- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIID ASSEMBLY BY
CC INTERACTING WITH BOTH TBP AND OTHER TAF, AS WELL AS SERVING TO
CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL
CC FOR PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES

```

```

CC DNA-BINDING ACTIVITY.
CC -1- SUBUNIT: TF2D IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED
CC FACTORS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: PHOSPHORYLATED BY CASEIN KINASE II IN VITRO.
CC -1- SIMILARITY: Contains 2 bromodomains.
CC -1- SIMILARITY: Contains 1 HMG box domain.
CC -1- SIMILARITY: TO DROSOPHILA TAFII-230. SOME TO S.POMBE TAFII-111
CC AND TO S.CEREVISIAE TAFI45.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90359; BAA14374.1; -
CC EMBL; X07024; CAA30073.1; ALT_SEQ.
CC PIR; A40262; A40262.
CC PDB; 1BOF; 07-JUN-00.
CC TRANSPAC; T02206; -
CC Genew; HGNC:11535; TAFI.
CC MIM; 313650; -
CC GO; GO:0005669; C:transcription factor TFIID complex; TAS.
CC GO; GO:0016251; F:general RNA polymerase II transcription fac. .; TAS.
CC GO; GO:0004672; F:protein kinase activity; TAS.
CC GO; GO:0000114; P:G1-specific transcription in mitotic cell c. .; TAS.
CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
CC GO; GO:0006461; P:protein complex assembly; TAS.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 2.
CC PROSITE; PSS00633; BROMODOMAIN_1; 2.
CC PROSITE; PSS0014; BROMODOMAIN_2; 2.
CC Bromodomain; Nuclear protein; DNA-binding; Cell cycle; Repeat;
KW Transcription regulation; Phosphorylation; 3D-structure.
FT DOMAIN 157 165 HMG BOX (POTENTIAL).
FT DNA BIND 1195 1273 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 1351 1358 BROMODOMAIN 1.
FT DOMAIN 1397 1467 BROMODOMAIN 2.
FT DOMAIN 1520 1590 BROMODOMAIN 1.
FT DOMAIN 1627 1872 ASP/GLU-RICH (ACIDIC TAIL).
SQ SEQUENCE 1872 AA; 212676 MW; 93BE3D181A72ABEB CRC64;

Query Match 75.5%; Score 40; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 1526 WPFHH 1530

RESULT 3
COAA_BACHD
ID COAA_BACHD STANDARD; PRT; 316 AA.
AC Q9KX7.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantothenate kinase (EC 2.7.1.33) (Pantothenic acid kinase).
GN COAA OR BH2875.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

```

```

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
CC -1- CATALYTIC ACTIVITY: ATP + pantothenate = ADP + D-4'-
CC phosphopantothenate.
CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC PANTOTHENATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U08189; AAA74739.1; -
DR PIR; S65669; S65669.
DR HSP; Q57366; I2U1.
DR InterPro; IPR006658; Bisc.
DR InterPro; IPR006657; Mol_dinuc_bind.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006655; Prok_Mboxred.
DR Pfam; PF00384; molybdopterin; 1.
DR Pfam; PF01568; molybdopterin; 1.
DR TIGRFAMs; TIGR00509; hScfam_1.
DR PROSITE; PS00551; MOLYBDOPROTEIN_PROK_1; FALSE_NEG.
DR PROSITE; PS00490; MOLYBDOPROTEIN_PROK_2; 1.
DR PROSITE; PS00932; MOLYBDOPROTEIN_PROK_3; FALSE_NEG.
DR Oxidoreductase; Molybdenum.
KW OXIDOREDUCTASE; MOLYBDENUM.
SQ SEQUENCE 744 AA; 80266 MW; 731A933E95358A19 CRC64;
Query Match 71.7%; Score 38; DB 1; Length 744;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 7
DB 401 PFHHH 406
RESULT 5
PIPI YEAST
ID PIPI YEAST STANDARD; PRT; 925 AA.
AC P40020;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymerase-interacting protein 1 (Factor interacting with REP).
GN PIPI OR FIR1 OR YER032W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Batstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RN Nature 387:78-81 (1997).
RL [3]
RN CHARACTERIZATION.
RP MEDLINE=97339480; PubMed=9196079;
RX Rusanak R., Pereira S., Platt T.;
RA "RNA binding analysis of yeast REF2 and its two-hybrid interaction
RT with a new gene product, Fir1.";
RL Gene Expr. 6:241-258 (1996).
CC -1- FUNCTION: Interacts with poly(A) polymerase and with REF2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U17262; AAB46625.1; --
 CC EMBL; U18778; AAB45855.1; --
 CC PIR; S50490; S50490.
 CC SGD; S0000834; F1R1.
 CC GO; GO:0006378; P:RNA polyadenylation; IGI.
 CC FT CONFLICT 663 R -> P (IN REF. 2).
 CC SQ SEQUENCE 925 AA; 104701 MW; 707D9839BE31322B CRC64;

Query Match 71.7%; Score 38; DB 1; Length 925;
 Best Local Similarity 71.4%; Pred. No. 75;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
 DB 179 WKPHNR 185

RESULT 6

HEX1-EMENI STANDARD; PRT; 221 AA.
 AC QP8K9;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Woronin body major protein.
 GN HEX1.
 OS Emericella nidulans (Aspergillus nidulans).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC Eurothiales; Trichocomaceae; Emericella.
 CC NCBI_TaxID=162425;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=20245872; PubMed=10783241;
 CC Jead G., Chua N.-H.;
 CC "A new self-assembled peroxisomal vesicle required for efficient
 CC resealing of the plasma membrane.";
 CC Nat. Cell Biol. 2:226-231(2000).
 CC -!- FUNCTION: Major protein in the hexagonal crystals of Woronin
 CC bodies, a peroxisomal vesicle that seal the septal pore in
 CC response to cellular damage.
 CC -!- SUBUNIT: Forms oligomers. Self-assemble into hexagonal rods (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Peroxisomal. Woronin bodies.
 CC -!- SIMILARITY: Belongs to the eIF-5A family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE239659; AAF67173.1; --
 CC InterPro; IPR001884; Eif5A_hypusine.
 CC Pfam; PF01287; eif-5a; 1.
 CC Peroxisome.
 CC FT SITE 219 221 MICROBODY TARGETING SIGNAL (POTENTIAL).
 CC SEQUENCE 221 AA; 25091 MW; 567634BFA7A62C93 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 221;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPHH 6
 DB 27 PPHH 31

RESULT 7

MRAY ARATH STANDARD; PRT; 236 AA.
 AC O49730;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phospho-N-acetylmuramoyl-pentapeptide-transferase homolog.
 GN A4G8270 OR T9A21.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=cv. Columbia;
 CC MEDLINE=20083488; PubMed=10617198;
 CC Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 CC Pohl T., Duesterhoef A., Stiekema W., Entian K.-D., Terry N.,
 CC Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 CC Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 CC Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
 CC Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 CC Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
 CC Langham S.-A., McCullagh B., Bilham L., Robben J., Vandenbussche F.,
 CC Van der Schueren J., Grymonprez B., Chuang Y.-J., Bastiaens I., Aert R.,
 CC Braeken M., Weltjens I., Voet M., Bastiaens I., Hilbert H., Braun M.,
 CC Holzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 CC Weitzer E., Brandt A., Peters S., van Staveren M., Dirksen W.,
 CC Moolman P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,
 CC Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 CC De Keyser A., Buyssehaert C., Gielen J., Villarroel R., De Clercq R.,
 CC Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 CC Clark L., Doggett J., Hall S., Kay M., Lennard N., McLaughlin K.,
 CC Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 CC Borkova D., Bloeker H., Scharif M., Grimm M., Loehner T.-H.,
 CC Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 CC Gabel C., Fuchs M., Fartmann B., Grandenath K., Dauner D., Herzl A.,
 CC Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
 CC Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 CC Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 CC Chedori F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 CC Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 CC Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 CC Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 CC Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 CC Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 CC Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 CC Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
 CC Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 CC Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 CC Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., L.,
 CC Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 CC Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 CC Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 CC Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 CC Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 CC Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 CC Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 CC Chen E., Marra M., Martienssen R., McCombie W.R.,
 CC "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 CC thaliana";
 CC RT thaliana;
 CC Nature 402:769-777(1999).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
 CC subfamily.
 CC -!- CAUTION: IT IS PROBABLE THAT THIS IS ONLY A FRAGMENT OF THE
 CC COMPLETE PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL021713; CAAL6799.1; -
 DR PIR; T04929; T04929.
 DR InterPro; IPR000715; Glyco_trans_4.
 DR InterPro; IPR003524; PNAcPept_trans.
 DR Pfam; PF00953; Glycos_transf_47_1.
 DR PROSITE; PS01347; MRAY_1; 1.
 DR PROSITE; PS01348; MRAY_2; 1.
 KW Hypothetical protein; Transferase; Transmembrane.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 47 67 POTENTIAL.
 FT TRANSMEM 80 100 POTENTIAL.
 FT TRANSMEM 112 132 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 SQ SEQUENCE 236 AA; 24830 MW; D93F9FCA086E3D28 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 236;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFRHH 6
 |||||
 Db 198 PFRHH 202

RESULT 8
 COL7 ARATH STANDARD; PRT; 327 AA.
 AC Q9SYM2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Zinc finger protein constans-like 7.
 GN AT1G78600 OR T30F21.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA White O., Alonso J., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Buehler E., Chan A., Chao Q., Conway A.R., Creasy T.H., Dewar K.,
 RA Chung M.K., Conn L., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Dunn P., Egu P., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Militescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., van Aken S., Vaysberg M., Vyotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana";
 RL Nature 408:816-820(2000).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
 CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AC007260; AAD30576.1; -
 DR PIR; F96814; F96814.
 DR InterPro; IPR000315; Znf_Box.
 DR InterPro; IPR002926; Znf_Constans.
 DR Pfam; PF00643; zf-B_box; 2.
 DR ProDom; PD007661; Znf_Constans; 2.
 DR SMART; SM00336; BBOX; 2.
 DR PROSITE; PS01119; ZF-BBOX; 2.
 KW Zinc-finger; Nuclear protein; Repeat; Multigene family.
 FT ZN_FING 5 47 B BOX-TYPE 1.
 FT ZN_FING 57 99 B BOX-TYPE 2.
 FT DOMAIN 149 155 POLY-GLN.
 SQ SEQUENCE 327 AA; 35392 MW; C8683E444D0504D6 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 327;
 Best Local Similarity 75.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 WP--FHHH 6
 |||||
 Db 291 WPKEFH 298

RESULT 9
 MRAY STRP3 STANDARD; PRT; 336 AA.
 ID MRAY STRP3
 AC Q8K6C7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-MurNAc-pentapeptide phosphotransferase).
 GN MRAY OR SPYMB_1400 OR SPS0462.
 OS Streptococcus pyogenes (serotype M3).
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus;
 CC Streptococcus;
 ON NCBI_TaxID=198466;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206;
 RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RN [2]
 SEQUENCE FROM N.A.
 RC STRAIN=SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 RA Yamazaki K., Okada H., Okada S., Kawabata S., Yasunaga T., Hattori M.,
 RA Hayashi H., Hamada S.;
 RT "The genome of invasive Streptococcus pyogenes; a comparative analysis of S. pyogenes SSI-1, SP370 and MGAS8232";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: First step of the lipid cycle reactions in the biosynthesis of the cell wall peptidoglycan (By similarity).
 CC -!- CATALYTIC ACTIVITY: UDPMurAc(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala)-diphosphoundecaprenol.
 CC -!- PATHWAY: Peptidoglycan biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE014163; AAM80007.1; -;
 CC EMBL; AP005142; BAC63557.1; -;
 CC HAMAP; MF_00038; -; 1;
 CC InterPro; IPR000715; Glyco_trans_4;
 CC InterPro; IPR003524; PNACpept_trans;
 CC Pfam; PF00953; Glycos_transf_4; 1;
 CC TIGRFAMs; TIGR00445; mray; 1;
 CC PROSITE; PS01347; MRAY_1; FALSE_NEG;
 CC PROSITE; PS01348; MRAY_2; 1;
 CC Peptidoglycan synthesis; Cell division; Transmembrane;
 CC Complete proteome.
 CC TRANSMEM 4 26 POTENTIAL.
 CC TRANSMEM 47 69 POTENTIAL.
 CC TRANSMEM 79 98 POTENTIAL.
 CC TRANSMEM 119 138 POTENTIAL.
 CC TRANSMEM 143 165 POTENTIAL.
 CC TRANSMEM 178 197 POTENTIAL.
 CC TRANSMEM 202 219 POTENTIAL.
 CC TRANSMEM 226 248 POTENTIAL.
 CC TRANSMEM 252 274 POTENTIAL.
 CC TRANSMEM 316 335 POTENTIAL.
 CC SEQUENCE 336 AA; 36886 MW; 7C8D20407E7B92 CRC64;
 CC
 CC Query Match 69.8%; Score 37; DB 1; Length 336;
 CC Best Local Similarity 100.0%; Pred. No. 42;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 PFHHH 6
 CC Db 292 PFHHH 296
 CC
 CC RESULT 10
 CC ID MRAY_STRP8 STANDARD; PRT; 336 AA.
 CC AC Q8NZY2;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
 CC MurNAC-pentapeptide phosphotransferase).
 CC GN MRAY OR SPY18_1673.
 CC OS Streptococcus pyogenes (serotype M18).
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=186103;
 CC RN [1]_TaxID=186103;
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=MGAS8232 / Serotype M18;
 CC RX MEDLINE=21927593; PubMed=11917108;
 CC RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 CC Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 CC Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 CC Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 CC "Genome sequence and comparative microarray analysis of serotype M18
 CC group A Streptococcus strains associated with acute rheumatic fever
 CC outbreaks";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 CC RL FUNCTION: First step of the lipid cycle reactions in the
 CC biosynthesis of the cell wall peptidoglycan (By similarity).
 CC CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
 CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
 CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
 CC CC -!- PATHWAY: Peptidoglycan biosynthesis.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE010079; AAL98215.1; -;
 CC HAMAP; MF_00038; -; 1;
 CC InterPro; IPR000715; Glyco_trans_4;
 CC InterPro; IPR003524; PNACpept_trans;
 CC Pfam; PF00953; Glycos_transf_4; 1;
 CC TIGRFAMs; TIGR00445; mray; 1;
 CC PROSITE; PS01347; MRAY_1; FALSE_NEG;
 CC PROSITE; PS01348; MRAY_2; 1;
 CC Peptidoglycan synthesis; Cell division; Transmembrane;
 CC Complete proteome.
 CC TRANSMEM 4 26 POTENTIAL.
 CC TRANSMEM 47 69 POTENTIAL.
 CC TRANSMEM 79 98 POTENTIAL.
 CC TRANSMEM 119 138 POTENTIAL.
 CC TRANSMEM 143 165 POTENTIAL.
 CC TRANSMEM 178 197 POTENTIAL.
 CC TRANSMEM 202 219 POTENTIAL.
 CC TRANSMEM 226 248 POTENTIAL.
 CC TRANSMEM 252 274 POTENTIAL.
 CC TRANSMEM 316 335 POTENTIAL.
 CC SEQUENCE 336 AA; 37075 MW; 6BFD8E55D71F67E2 CRC64;
 CC
 CC Query Match 69.8%; Score 37; DB 1; Length 336;
 CC Best Local Similarity 100.0%; Pred. No. 42;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 PFHHH 6
 CC Db 292 PFHHH 296
 CC
 CC RESULT 11
 CC ID MRAY_STRPY STANDARD; PRT; 336 AA.
 CC AC Q99YK2;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
 CC MurNAC-pentapeptide phosphotransferase).
 CC GN MRAY OR SP1682.
 CC OS Streptococcus pyogenes.
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=1314;
 CC RN [1]_TaxID=1314;
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 CC RX MEDLINE=21192684; PubMed=11296296;
 CC RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 CC Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 CC Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 CC Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 CC "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 CC Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 CC RL FUNCTION: First step of the lipid cycle reactions in the
 CC biosynthesis of the cell wall peptidoglycan (By similarity).
 CC CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
 CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
 CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
 CC CC -!- PATHWAY: Peptidoglycan biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AEO06597; AAK34425.1; -
CC HAMAP: MF_00038; -; 1.
CC InterPro: IPR000715; Glyco trans 4.
CC InterPro: IPR003524; PNACpppt trans.
CC Pfam: PF00953; Glycos transf_4; 1.
CC TIGRFAMs: TIGR00445; mray; 1.
CC PROSITE: PS01347; MRAY_1; FALSE_NEG.
CC PROSITE: PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transmembrane;
KW Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 47 69 POTENTIAL.
FT TRANSMEM 79 98 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT TRANSMEM 143 165 POTENTIAL.
FT TRANSMEM 178 197 POTENTIAL.
FT TRANSMEM 202 219 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 252 274 POTENTIAL.
FT TRANSMEM 316 335 POTENTIAL.
SQ SEQUENCE 336 AA; 36924 MW; 093C5B79CC0510CD CRC64;
Query Match 69.8%; Score 37; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
DB 292 PFHHH 296
RESULT 12
MRAY_AQAE STANDARD; PRT; 359 AA.
AC O66465;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR AQ_053.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AEO0671; AAC06418.1; -
CC PIR: P70304; F70304.1;
CC HAMAP: MF_00038; -; 1.
CC InterPro: IPR000715; Glyco trans 4.
CC InterPro: IPR003524; PNACpppt trans.
CC Pfam: PF00953; Glycos transf_4; 1.
CC TIGRFAMs: TIGR00445; mray; 1.
CC PROSITE: PS01347; MRAY_1; 1.
CC PROSITE: PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transmembrane;
KW Complete proteome.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 198 218 POTENTIAL.
FT TRANSMEM 234 254 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 287 307 POTENTIAL.
FT TRANSMEM 336 356 POTENTIAL.
SQ SEQUENCE 359 AA; 40341 MW; B2D6292EA6A16D4 CRC64;
Query Match 69.8%; Score 37; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
DB 322 PFHHH 326
RESULT 13
MRAY_MYCLE STANDARD; PRT; 359 AA.
AC O69555;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR MURX OR MLO911 OR MLCB268.05C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RL "Massive gene decay in the leprosy bacillus";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE

CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -1- PATHWAY: Peptidoglycan biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. May

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AL022602; CAA18671.1; -

CC EMBL; AL583920; CAC31292.1; -

CC F1R; A87023; A87023.

CC Leproma; MF_00038; -; 1.

CC InterPro; IPR000715; Glyco trans 4.

CC InterPro; IPR003524; PNAcPpPpt trans.

CC Pfam; PF00953; Glycos transf 4; 1.

CC TIGR; TIGR00445; mray; 1.

CC PROSITE; PS01347; MRAY_1; 1.

CC PROSITE; PS01348; MRAY_2; 1.

CC Peptidoglycan synthesis; Cell division; Transmembrane;

CC Complete proteome.

CC TRANSMEM 3 23 POTENTIAL.

CC TRANSMEM 55 75 POTENTIAL.

CC TRANSMEM 80 100 POTENTIAL.

CC TRANSMEM 117 137 POTENTIAL.

CC TRANSMEM 156 176 POTENTIAL.

CC TRANSMEM 187 207 POTENTIAL.

CC TRANSMEM 231 251 POTENTIAL.

CC TRANSMEM 255 275 POTENTIAL.

CC TRANSMEM 280 300 POTENTIAL.

CC TRANSMEM 334 354 POTENTIAL.

CC SEQUENCE 359 AA; 37773 MW; 15554A306A3B4429 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 359;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHH 6

Db 315 PFHHH 319

RESULT 14

MRAY MYCTU STANDARD; PRT; 359 AA.

AC O06221; (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

DE MurNAc-pentapeptide phosphotransferase).

GN MRAY OR MURX OR RV2156C OR M2215 OR MTCY270.12.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekalia F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

RA Sultun J.S., Taylor K., Whitehead S., Barrell B.G.,

RA "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence."

RL Nature 393:537-544(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Emdinova M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W.

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE

CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -1- PATHWAY: Peptidoglycan biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. May

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z95388; CAB08671.1; -

CC EMBL; AB007068; AAK46499.1; -

CC F1R; H70579; H70579.

CC TIGR; MT2215; -

CC Tuberculist; RV2156c; -

CC HAMAP; MF_00038; -; 1.

CC InterPro; IPR000715; Glyco trans 4.

CC InterPro; IPR003524; PNAcPpPpt trans.

CC Pfam; PF00953; Glycos transf 4; 1.

CC TIGR; TIGR00445; mray; 1.

CC PROSITE; PS01347; MRAY_1; 1.

CC PROSITE; PS01348; MRAY_2; 1.

CC Peptidoglycan synthesis; Cell division; Transmembrane;

CC Complete proteome.

CC TRANSMEM 3 23 POTENTIAL.

CC TRANSMEM 55 75 POTENTIAL.

CC TRANSMEM 80 100 POTENTIAL.

CC TRANSMEM 117 137 POTENTIAL.

CC TRANSMEM 156 176 POTENTIAL.

CC TRANSMEM 187 207 POTENTIAL.

CC TRANSMEM 231 251 POTENTIAL.

CC TRANSMEM 255 275 POTENTIAL.

CC TRANSMEM 280 300 POTENTIAL.

CC TRANSMEM 334 354 POTENTIAL.

CC SEQUENCE 359 AA; 37713 MW; 27E979BC77BA28C6 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 359;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHH 6

Db 315 PFHHH 319

RESULT 15

NY4R_HUMAN

KW	G-protein coupled receptor; Transmembrane; Glycoprotein;
KW	phosphorylation; Lipoprotein; Palmitate.
FT	DOMAIN 1 40
FT	TRANSMEM 41 53
FT	DOMAIN 64 74
FT	TRANSMEM 75 95
FT	DOMAIN 96 115
FT	TRANSMEM 116 137
FT	DOMAIN 138 157
FT	TRANSMEM 158 178
FT	DOMAIN 179 213
FT	TRANSMEM 214 235
FT	DOMAIN 236 265
FT	TRANSMEM 266 288
FT	DOMAIN 289 301
FT	TRANSMEM 302 325
FT	DOMAIN 326 375
FT	CARBOHYD 2 2
FT	CARBOHYD 19 19
FT	CARBOHYD 29 29
FT	CARBOHYD 187 187
FT	DISULFID 114 201
FT	LIPID 340 340
FT	CONFLICT 99 99
FT	SEQUENCE 375 AA; 42195 MW; 0081DBA51436BE22 CRC64;
QY	Query Match 59.8%; Score 37; DB 1; Length 375;
Db	Best Local Similarity 71.4%; Pred. No. 47; Indels 0; Gaps 0;
	Matches 5; Conservative 0; Mismatches 2
Qy	1 WPFHHR 7
Db	205 WPLAHR 211

Search completed: February 11, 2004, 17:04:13
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-9

Perfect score: 53

Sequence: 1 WPFHHH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	310	2 T46035	AtMYB84 - Arabidop
2	42	79.2	423	2 E64157	hypothetical prote
3	40	75.5	1865	1 I48155	transcription init
4	40	75.5	1893	1 A40262	transcription init
5	39	73.6	316	2 C84009	pantothenate kinas
6	39	73.6	750	2 T04980	hypothetical prote
7	38.5	72.6	750	2 T38435	coiled coil protei
8	38	71.7	284	2 A70976	hypothetical prote
9	38	71.7	293	2 S64925	hypothetical prote
10	38	71.7	316	2 G82204	transcription regu
11	38	71.7	674	2 T21217	hypothetical prote
12	38	71.7	744	2 S65669	biotin sulfoxide r
13	38	71.7	753	2 AH0097	probable biotin su
14	38	71.7	766	2 E70916	probable biot prot
15	38	71.7	925	2 S50490	hypothetical prote
16	37.5	70.8	79	2 T30118	hypothetical prote
17	37	69.8	236	2 T04929	hypothetical prote
18	37	69.8	313	2 F85205	hypothetical prote
19	37	69.8	327	2 F96814	hypothetical prote
20	37	69.8	359	2 F70304	phospho-N-acetylmu
21	37	69.8	359	2 A87023	hypothetical prote
22	37	69.8	359	2 H70579	probable murx prot
23	37	69.8	375	2 G02300	pancreatic polypep
24	37	69.8	375	2 I33182	neuropeptide Y/pep
25	37	69.8	395	2 S39659	dlcB protein - Bac
26	37	69.8	520	2 F86401	protein f22C5.11 [
27	37	69.8	557	2 AC2137	ABC transporter At
28	37	69.8	611	2 G96032	probable thiamine
29	37	69.8	710	2 T47610	hypothetical prote

ALIGNMENTS

RESULT 1

T46035
AtMYB84 - Arabidopsis thaliana
N:Alternate names: protein T16K5.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000
C:Accession: T46035
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23015
A:Accession: T46035
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <RIB>
A:Cross-references: EMBL:AL132965
A:Experimental source: Cultivar Columbia; BAC clone T16K5
C:Genetics:
A:Map position: 3
A:Introns: 46/1; 89/2
A:Notes: T16K5.40
C:Superfamily: Arabidopsis myb-related protein Y13; myb DNA-binding repeat homology

Query Match 79.2%; Score 42; DB 2; Length 310;
Best Local Similarity 83.3%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6

DB 163 WPLHHH 168

RESULT 2

E64157
hypothetical protein HI0723 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Sep-1998
C:Accession: E64157
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, R.; Fleischmann, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, G.; Gocayne, J.D.; Scott, J.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: E64157
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-423 <TIGR>
A:Cross-references: GB:U32755; GB:L42023; NID:g1573722; PID:g1573726; TIGR:HI0723
A:Note: best homolog was a hypothetical protein from Escherichia coli
C:Genetics:
A:Start codon: GTG

probable helicase
hypothetical prote
hypothetical prote
hemoglobin alpha c
hypothetical prote
transcription regu
probable oxidoredu
probable N-acyl-L-
nating pheromone r
3',5'-cyclic-nucle
unknown protein T1
T-cell receptor be
cholesterol salph
hypothetical prote
T-cell receptor be
transcription fact

30 37 69.8 1040 2 T39936
31 37 69.8 1091 2 T48444
32 36.5 68.9 332 2 D70716
33 36 67.9 101 2 T49585
34 36 67.9 243 2 T38519
35 36 67.9 297 2 B84088
36 36 67.9 339 2 C95358
37 36 67.9 332 2 C75331
38 36 67.9 404 2 T50335
39 36 67.9 526 2 S67272
40 36 67.9 676 2 A96722
41 35 66.0 140 2 I46880
42 35 66.0 254 2 A49169
43 35 66.0 300 2 T33232
44 35 66.0 319 1 RWRBB
45 35 66.0 369 1 S23248

C:Superfamily: potassium uptake protein trkG

Query Match 79.2%; Score 42; DB 2; Length 423;
Best Local Similarity 71.4%; Pred. No. 9.5;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 WPFHHH 7
DB 55 WPCCHHK 61

RESULT 3
I48155
transcription initiation factor IID 250K chain splice form 2 - golden hamster
N:Alternate names: 210K nuclear DNA-binding cell cycle gene 1 protein (CCG1); CCG1; TATA
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 15-Oct-1999
C:Accession: I48155
R:Hayashida, T.; Sekiguchi, T.; Noguchi, E.; Sunamoto, H.; Ohba, T.; Nishimoto, T.
Gene 141, 267-270, 1994
A:Title: The CCG1/TAFII250 gene is mutated in thermosensitive G1 mutants of the BHK21 cell line
A:Reference number: I48155; MUID:94215915; PMID:8163200
A:Accession: I48155
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1865 <RES>
A:Cross-references: GB:D26114; NID:9439485; PIDN:BAA05110.1; PID:9474971
C:Superfamily: transcription initiation factor IID 250K chain; bromodomain homology; HMG
C:Keywords: alternative splicing; cell cycle control; DNA binding; duplication; phospho
F:1190-1269/Domain: HMG box homology <HMG1>
F:1346-1353/Region: nuclear location signal
F:1400-1455/Domain: bromodomain homology <BR01>
F:1523-1576/Domain: bromodomain homology <BR02>
F:131,171,174,175,182,183/Binding site: phosphate (Ser) (covalent) (by casein kinase II)
F:652,1028,1658/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase)
F:994,1335/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status F
F:1355,1374/Binding site: phosphate (Thr) (covalent) (by calmodulin-dependent kinase) #s

Query Match 75.5%; Score 40; DB 1; Length 1865;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
DB 1521 WPFHH 1525

RESULT 4
A40262
transcription initiation factor IID 250K chain splice form 1 - human
N:Alternate names: 210K nuclear DNA-binding cell cycle gene 1 protein (CCG1); TATA-bind
N:Contains: transcription initiation factor IID 250K chain splice form 2
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A40262; S03005; S00830; S32352; S32353
R:Sekiguchi, T.; Nohiro, Y.; Nakamura, Y.; Hisamoto, N.; Nishimoto, T.
Mol. Cell. Biol. 11, 3317-3325, 1991
A:Title: The human CCG1 gene, essential for progression of the G-1 phase, encodes a 210-
A:Reference number: A40262; MUID:91246200; PMID:2038334
A:Accession: A40262
A:Molecule type: mRNA
A:Residues: 1-177,199-1893 <SEK>
A:Cross-references: GB:D90359; NID:9559319; PIDN:BAAL4374.1; PID:g219528
A:Note: nucleotide sequence not complete
R:Sekiguchi, T.; Miyata, T.; Nishimoto, T.
submitted to the EMBL Data Library, February 1989
A:Reference number: S03005
A:Accession: S03005
A:Molecule type: mRNA
A:Residues: 'MIR', 60-177,199-1604, 'DNECSKANDIVCLIQYSSQIEELRP' <SE5>
A:Cross-references: EMBL:X07024; NID:929732; PIDN:CAA30073.1; PID:g29733
A:Note: this sequence has been revised in reference A40262
R:Sekiguchi, T.; Miyata, T.; Nishimoto, T.

EMBO J. 7, 1683-1687, 1988
A:Title: Molecular cloning of the cDNA of human X chromosomal gene (CCG1) which complem
A:Reference number: S00830; MUID:89005056; PMID:3169001
A:Accession: S00830
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1372-1379,1386-1604, 'DNE', 'CSSKANDIVCLIQYSSQIEELRP' <SE6>
A:Cross-references: EMBL:X07024
R:Ruppert, S.; Wang, E.H.; Tjian, R.
Nature 362, 175-179, 1993
A:Title: Cloning and expression of human TAF(II)250: a TBP-associated factor implicated
A:Reference number: S32352; MUID:93196704; PMID:7680771
A:Accession: S32352

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 178-198 <RUP>
R:Hisatake, K.; Hasegawa, S.; Takada, R.; Nakatani, Y.; Horikoshi, M.; Roeder, R.G.
Nature 362, 179-181, 1993
A:Title: The p250 subunit of native TATA box-binding factor TFIID is the cell-cycle reg
A:Reference number: S32353; MUID:93196705; PMID:8450888
A:Accession: S32353
A:Status: preliminary
A:Molecule type: Protein
A:Residues: 'P', 587-595, 1009-1022,1351-1355,1357-1360 <HIS>
A:Note: 1351-Val, 1353-Lys, and 1354-Glu were also found
C:Genetics:

A:Gene: GDB:TAF2A; CCG1; CCGS; NSCL2; TAFII250
A:Cross-references: GDB:120573; OMIM:313650
A:Map position: Xq13.1-Xq13.1
C:Superfamily: transcription initiation factor IID 250K chain; bromodomain homology; HM
C:Keywords: alternative splicing; cell cycle control; DNA binding; duplication; phospho
F:1-1893/Product: transcription initiation factor IID 250K chain splice form 1 #status
F:1-177,199-1893/Product: transcription initiation factor IID 250K chain splice form 2
F:1216-1295/Domain: HMG box homology <HMG1>
F:1372-1379/Region: nuclear location signal
F:1426-1481/Domain: bromodomain homology <BR01>
F:1549-1604/Domain: bromodomain homology <BR02>
F:137,174,175,182,183/Binding site: phosphate (Ser) (covalent) (by casein kinase II)
F:678,1054,1684/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase)
F:1020,1361/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status
F:1381,1400/Binding site: phosphate (Thr) (covalent) (by calmodulin-dependent kinase) #

Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
DB 1547 WPFHH 1551

RESULT 5
C84009
pantothenate kinase BH2875 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: C84009
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84009
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-316 <STO>
A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA806594.1; GSPDB:GNO
A:Experimental source: strain C-125
C:Genetics:
C:Gene: BH2875
C:Superfamily: pantothenate kinase

Query Match 73.6%; Score 39; DB 2; Length 316;
Best Local Similarity 83.3%; Pred. No. 22;

Query Match 71.7%; Score 38; DB 2; Length 316;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 212 WIFHHH 217

RESULT 11
T21217
hypothetical protein F21G4.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21217
R:Mortimore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19392
A:Accession: T21217
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-674 <WIL>
A:Cross-references: EMBL:Z81016; PIDN:CA802665.1; GSPDB:GN00028; CESP:F21G4.1
A:Experimental source: clone F21G4
C:Genetics:
A:Gene: CESP:F21G4.1
A:Map position: X
A:Introns: 58/3, 149/1, 181/3, 213/3, 286/1, 327/2, 457/2, 509/3, 608/1

Query Match 71.7%; Score 38; DB 2; Length 674;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 7
Db 458 WHFYHR 464

RESULT 12
S65669
biotin sulfoxide reductase (BC 1.8.4.-) - *Rhodobacter sphaeroides*
C:Species: *Rhodobacter sphaeroides*
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S65669
R:Pollock, V.V.; Barber, M.J.
Arch. Biochem. Biophys. 318, 322-332, 1995
A:Title: Molecular cloning and expression of biotin sulfoxide reductase from *Rhodobacter sphaeroides*
A:Reference number: S65669; MUID:95251380; PMID:7733660
A:Accession: S65669
A:Molecule type: DNA
A:Residues: 1-744 <POL>
A:Cross-references: EMBL:U08189; NID:G953223; PIDN:AAA74739.1; PID:G953224
A:Experimental source: strain Forma sp. denitrificans
A:Note: the authors translated the initiation codon GTG for residue 1 as Val
C:Genetics:
A:Start codon: GTG
C:Superfamily: trimethylamine-N-oxide reductase
C:Keywords: molybdenum; oxidoreductase

Query Match 71.7%; Score 38; DB 2; Length 744;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPFHHR 7
Db 401 PPFHHQ 406

RESULT 13
AH0097
probable biotin sulfoxide reductase2 bisc [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0097
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0097
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-753 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA89643.1; PID:gl5978871; GSPDB:GN00175
C:Genetics:
A:Gene: bisc
C:Superfamily: trimethylamine-N-oxide reductase

Query Match 71.7%; Score 38; DB 2; Length 753;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPFHHR 7
Db 425 PPFHHQ 430

RESULT 14
E70916
probable bisc protein - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: E70916
R:Col, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70916
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-766 <COL>
A:Cross-references: GB:Z95844; GB:AL123456; NID:G3250713; PIDN:CA809264.1; PID:G2131054
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: bisc
C:Superfamily: trimethylamine-N-oxide reductase

Query Match 71.7%; Score 38; DB 2; Length 766;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPFHHR 7
Db 427 PPFHHQ 432

RESULT 15
S50490
hypothetical protein YER032w - yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C>Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001
C:Accession: S50490
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9537, 9581, 9495, 9867, and lambda
A:Reference number: S50433
A:Accession: S50490
A:Molecule type: DNA
A:Residues: 1-925 <DIE>
A:Cross-references: EMBL:U18778; NID:G603592; PIDN:AA64565.1; PID:G603624; MIPS:YER032;
C:Genetics:

A:Gene: SGD:FIR1
A:Cross-references: SGD:S0000834; MIPS:YER032w
A:Map position: 5R

Query Match 71.7%; Score 38; DB 2; Length 925;
Best Local Similarity 71.4%; Pred. NO. 97;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
DB 179 WKFHNR 185

Search completed: February 11, 2004, 17:11:46
Job time : 13.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-9

Perfect score: 53

Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 20938283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	7	10	US-09-901-187B-9
2	45	84.9	27	12	US-10-029-386-30005
3	42	79.2	309	12	US-10-374-780A-2268
4	40	75.5	111	12	US-10-278-536-120
5	40	75.5	111	12	US-10-209-201C-17
6	40	75.5	111	12	US-10-209-201C-18
7	40	75.5	128	15	US-10-012-542-465
8	40	75.5	165	15	US-10-012-542-464
9	40	75.5	898	15	US-10-043-487-277
10	39	73.6	16	10	US-09-948-018-25
11	38.5	72.6	115	9	US-09-989-722-95
12	38.5	72.6	115	9	US-09-989-723-95
13	38.5	72.6	115	9	US-09-989-279-95
14	38.5	72.6	115	9	US-09-989-727-95
15	38.5	72.6	115	10	US-09-989-731-95

16	38.5	72.6	115	10	US-09-989-732-95	Sequence 95, Appl
17	38.5	72.6	115	10	US-09-991-073-95	Sequence 95, Appl
18	38.5	72.6	115	10	US-09-990-442-95	Sequence 95, Appl
19	38.5	72.6	115	10	US-09-991-163-95	Sequence 95, Appl
20	38.5	72.6	115	10	US-09-993-604-95	Sequence 95, Appl
21	38.5	72.6	115	10	US-09-990-456-95	Sequence 95, Appl
22	38.5	72.6	115	10	US-09-989-721-95	Sequence 95, Appl
23	38.5	72.6	115	10	US-09-992-598-95	Sequence 95, Appl
24	38.5	72.6	115	10	US-09-989-293A-95	Sequence 95, Appl
25	38.5	72.6	115	10	US-09-989-735-95	Sequence 95, Appl
26	38.5	72.6	115	10	US-09-990-444-95	Sequence 95, Appl
27	38.5	72.6	115	10	US-09-991-181-95	Sequence 95, Appl
28	38.5	72.6	115	10	US-09-989-730-95	Sequence 95, Appl
29	38.5	72.6	115	10	US-09-990-436-95	Sequence 95, Appl
30	38.5	72.6	115	10	US-09-993-687-95	Sequence 95, Appl
31	38.5	72.6	115	11	US-09-989-734-95	Sequence 95, Appl
32	38.5	72.6	115	11	US-09-997-653-95	Sequence 95, Appl
33	38.5	72.6	115	11	US-09-993-667-95	Sequence 95, Appl
34	38.5	72.6	115	11	US-09-997-428-95	Sequence 95, Appl
35	38.5	72.6	115	11	US-09-997-666-95	Sequence 95, Appl
36	38.5	72.6	115	11	US-09-990-438-95	Sequence 95, Appl
37	38.5	72.6	115	11	US-09-990-562-95	Sequence 95, Appl
38	38.5	72.6	115	11	US-09-990-711-95	Sequence 95, Appl
39	38.5	72.6	115	11	US-09-989-726-95	Sequence 95, Appl
40	38.5	72.6	115	11	US-09-988-156-95	Sequence 95, Appl
41	38.5	72.6	115	11	US-09-990-437-95	Sequence 95, Appl
42	38.5	72.6	115	11	US-09-991-157-95	Sequence 95, Appl
43	38.5	72.6	115	11	US-09-997-514-95	Sequence 95, Appl
44	38.5	72.6	115	11	US-09-997-573-95	Sequence 95, Appl
45	38.5	72.6	115	11	US-09-991-172-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1

US-09-901-187B-9

; Sequence 9, Application US/09901187B

; Patent No. US20020151464A1

; GENERAL INFORMATION:

; APPLICANT: Panacea Pharmaceuticals, Inc.

; APPLICANT: Wolozin, Benjamin

; APPLICANT: Ostretova-Golts, Natalie

; APPLICANT: Lebowitz, Micheal S.

; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment

; TITLE OF INVENTION: Alpha-Synuclein Diseases

; FILE REFERENCE: PAN01/00205

; CURRENT FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/217,319

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/279,199

; PRIOR FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-901-187B-9

Query Match 100.0%; Score 53; DB 10; Length 7;

Best Local Similarity 100.0%; Pred. No. 7e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHHR 7

Db 1 WPFHHR 7

RESULT 2

US-10-029-386-30005

; Sequence 30005, Application US/10029386

Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AECMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30005
LENGTH: 27
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ALL36530.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.6
US-10-029-386-30005
Query Match 84.9%; Score 45; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 3.5;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHHH 6
DB 5 WPFHHH 10
RESULT 3
US-10-374-780A-2268
Sequence 2268, Application US/10374780A
Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2268
LENGTH: 309
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G234
US-10-374-780A-2268
Query Match 79.2%; Score 42; DB 12; Length 309;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WPFHHH 6
DB 163 WPLHHH 168
RESULT 4
US-10-278-536-120
Sequence 120, Application US/10278536
Publication No. US20030131386A1
GENERAL INFORMATION:
APPLICANT: Samaha, Raymond
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Keddie, James
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Adam, Luc
APPLICANT: Broun, Pierre
TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-011
CURRENT APPLICATION NUMBER: US/10/278,536
CURRENT FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 120
LENGTH: 309
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G234
US-10-278-536-120
Query Match 79.2%; Score 42; DB 16; Length 309;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WPFHHH 6
DB 163 WPLHHH 168
RESULT 5
US-10-209-201C-17
Sequence 17, Application US/10209201C
Publication No. US20040009613A1
GENERAL INFORMATION:
APPLICANT: Verdin, Eric
APPLICANT: Bruland, Joan
APPLICANT: Ott, Melanie

```

; APPLICANT: Zhou, Ming-Wing
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 17
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-17

Query Match      75.5%; Score 40; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WPFHH 5
Db      29 WPFHH 33

RESULT 6
US-10-209-201C-18
; Sequence 18; Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Mesocricetus auratus
US-10-209-201C-18

Query Match      75.5%; Score 40; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WPFHH 5
Db      29 WPFHH 33

RESULT 7
US-10-012-542-465
; Sequence 465; Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 09/461,325

```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 465
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-465

Query Match      75.5%; Score 40; DB 15; Length 128;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHHR 7
Db      50 WPFHHR 56

RESULT 8
US-10-012-542-464
; Sequence 464; Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 464
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-464

Query Match      75.5%; Score 40; DB 15; Length 165;
Best Local Similarity 71.4%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHHR 7
Db      87 WPFHHR 93

RESULT 9
US-10-043-487-277

```

```

; Sequence 277, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: BIERRE, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 277
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-277

Query Match          75.5%; Score 40; DB 15; Length 898;
Best Local Similarity 100.0%; Pred. No. 3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 552 WPFHH 556

RESULT 10
US-09-948-018-25
; Sequence 25, Application US/09948018
; Patent No. US20020150977A1
; GENERAL INFORMATION:
; APPLICANT: Theill et al
; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37677
; CURRENT APPLICATION NUMBER: US/09/948,018
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,191
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-018-25

Query Match          73.6%; Score 39; DB 10; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHH 6
Db 8 WPNHH 13

RESULT 11
US-09-989-722-95
; Sequence 95, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austin L.
; APPLICANT: KJavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoli, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C63
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09

```

; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535

; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 9; Length 115;
 Best Local Similarity 43.8%; Pred. No. 99;
 Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 WP-----FHHR 7
 DB 59 WPFRRRGLGIFHHR 74

RESULT 12
 US-09-989-723-95
 ; Sequence 95, Application US/09989723
 ; Patent No. US20020072092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tamas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2730P1C62
 CURRENT APPLICATION NUMBER: US/09/989,723
 CURRENT FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 9; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 WP-----FHHR 7
Db 59 WPPRRGHGIGFHHR 74

RESULT 13

US-09-989-279-95
Sequence 95, Application US/09989279
Patent No. US20020072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Saton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC56
CURRENT APPLICATION NUMBER: US/09/989,279
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11

```

RESULT 14
US - 95-989-727-95
Sequence 98, Application US/09898727
Patent No. US20020072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Petrakaki, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and T
TITLE OF INVENTION: Acids Encoding
FILE REFERENCE: F2730PIC65
CURRENT APPLICATION NUMBER: US/09/09/989727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/04978
PRIOR FILING DATE: 1997-06-16

```

1	PRIOR FILING DATE: 1998-06-16	
2	PRIOR APPLICATION NUMBER: 60/089512	
3	PRIOR FILING DATE: 1998-06-16	
4	PRIOR APPLICATION NUMBER: 60/089514	
5	PRIOR FILING DATE: 1998-06-16	
6	PRIOR APPLICATION NUMBER: 60/089532	
7	PRIOR FILING DATE: 1998-06-17	
8	PRIOR APPLICATION NUMBER: 60/089538	
9	PRIOR FILING DATE: 1998-06-17	
10	PRIOR APPLICATION NUMBER: 60/089598	
11	PRIOR FILING DATE: 1998-06-17	
12	PRIOR APPLICATION NUMBER: 60/089599	
13	PRIOR FILING DATE: 1998-06-17	
14	PRIOR APPLICATION NUMBER: 60/089600	
15	PRIOR FILING DATE: 1998-06-17	
16	PRIOR APPLICATION NUMBER: 60/089653	
17	PRIOR FILING DATE: 1998-06-17	
18	PRIOR APPLICATION NUMBER: 60/089801	
19	PRIOR FILING DATE: 1998-06-18	
20	PRIOR APPLICATION NUMBER: 60/089907	
21	PRIOR FILING DATE: 1998-06-18	
22	PRIOR APPLICATION NUMBER: 60/089908	
23	PRIOR FILING DATE: 1998-06-18	
24	PRIOR APPLICATION NUMBER: 60/089947	
25	PRIOR FILING DATE: 1998-06-19	
26	PRIOR APPLICATION NUMBER: 60/089948	
27	PRIOR FILING DATE: 1998-06-19	
28	PRIOR APPLICATION NUMBER: 60/089952	
29	PRIOR FILING DATE: 1998-06-19	
30	PRIOR APPLICATION NUMBER: 60/090246	
31	PRIOR FILING DATE: 1998-06-22	
32	PRIOR APPLICATION NUMBER: 60/090252	
33	PRIOR FILING DATE: 1998-06-22	
34	PRIOR APPLICATION NUMBER: 60/090254	
35	PRIOR FILING DATE: 1998-06-22	
36	PRIOR APPLICATION NUMBER: 60/090349	
37	PRIOR FILING DATE: 1998-06-23	
38	PRIOR APPLICATION NUMBER: 60/090355	
39	PRIOR FILING DATE: 1998-06-23	
40	PRIOR APPLICATION NUMBER: 60/090429	
41	PRIOR FILING DATE: 1998-06-24	
42	PRIOR APPLICATION NUMBER: 60/090431	
43	PRIOR FILING DATE: 1998-06-24	
44	PRIOR APPLICATION NUMBER: 60/090435	
45	PRIOR FILING DATE: 1998-06-24	
46	PRIOR APPLICATION NUMBER: 60/090444	
47	PRIOR FILING DATE: 1998-06-24	
48	PRIOR APPLICATION NUMBER: 60/090445	
49	PRIOR FILING DATE: 1998-06-24	
50	PRIOR APPLICATION NUMBER: 60/090472	
51	PRIOR FILING DATE: 1998-06-24	
52	PRIOR APPLICATION NUMBER: 60/090535	
53	PRIOR FILING DATE: 1998-06-24	
54	PRIOR APPLICATION NUMBER: 60/090540	
55	PRIOR FILING DATE: 1998-06-24	
56	PRIOR APPLICATION NUMBER: 60/090542	
57	PRIOR FILING DATE: 1998-06-24	
58	PRIOR APPLICATION NUMBER: 60/090557	
59	PRIOR FILING DATE: 1998-06-24	
60	PRIOR APPLICATION NUMBER: 60/090676	
61	PRIOR FILING DATE: 1998-06-25	
62	PRIOR APPLICATION NUMBER: 60/090678	
63	PRIOR FILING DATE: 1998-06-25	
64	PRIOR APPLICATION NUMBER: 60/090690	
65	PRIOR FILING DATE: 1998-06-25	
66	PRIOR APPLICATION NUMBER: 60/090694	
67	PRIOR FILING DATE: 1998-06-25	
68	PRIOR APPLICATION NUMBER: 60/090695	
69	PRIOR FILING DATE: 1998-06-25	
70	PRIOR APPLICATION NUMBER: 60/090696	
71	PRIOR FILING DATE: 1998-06-25	
72	PRIOR APPLICATION NUMBER: 60/090862	
73	PRIOR FILING DATE: 1998-06-26	

PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 9; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 9; Gaps 1;

QY 1 WP-----FHHR 7
DB 59 WFFRRRGLGIFHHHR 74

RESULT 15
US-09-989-731-95
Sequence 95, Application US/09989731
Patent No. US20020103125A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544

; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 10; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 WP-----FHHHR 7
|||
Db 59 WFFRRCHLGIFFHHR 74

Search completed: February 11, 2004, 17:54:12
Job time : 25.6667 secs

Sequence 7, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 26036, A
Sequence 20452, A
Sequence 32813, A
Sequence 24508, A
Sequence 29635, A
Sequence 30587, A
Sequence 4, Appli
Sequence 12, Appli
Sequence 58, Appli
Sequence 58, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 56, Appli

28 35 66.0 241 4 US-09-709-791-7
29 35 66.0 254 1 US-07-795-859B-6
30 35 66.0 254 1 US-08-457-616-6
31 35 66.0 262 4 US-09-252-991A-26036
32 35 66.0 368 4 US-09-252-991A-20452
33 35 66.0 427 4 US-09-252-991A-32813
34 35 66.0 481 4 US-09-252-991A-24508
35 35 66.0 525 4 US-09-252-991A-29635
36 35 66.0 694 4 US-09-252-991A-30587
37 35 66.0 919 2 US-08-788-674-4
38 35 66.0 942 4 US-09-171-461-12
39 35 66.0 951 3 US-08-816-346-58
40 35 66.0 951 3 US-09-335-411-58
41 35 66.0 952 3 US-08-788-674-5
42 35 66.0 952 3 US-08-816-346-4
43 35 66.0 952 3 US-09-335-411-4
44 35 66.0 967 3 US-08-816-346-56
45 35 66.0 967 3 US-09-335-411-56

ALIGNMENTS

RESULT 1
US-09-252-991A-22036
; Sequence 22036, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22036
; LENGTH: 580
; TYPE: PRI
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22036

Query Match 84.9%; Score 45; DB 4; Length 580;
Best Local Similarity 85.7%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7

Db 294 WPAHHR 300

RESULT 2
US-08-227-536-4
; Sequence 4, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-9

Perfect score: 53

Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	84.9	580	US-09-252-991A-22036	Sequence 22036, A
2	40	75.5	65	US-08-227-536-4	Sequence 4, Appli
3	40	75.5	65	PCT-US95-04682-4	Sequence 4, Appli
4	40	75.5	128	US-09-461-325-465	Sequence 465, App
5	40	75.5	165	US-09-461-325-464	Sequence 464, App
6	40	75.5	1872	US-08-188-582-14	Sequence 14, Appl
7	40	75.5	1872	US-08-646-715-14	Sequence 14, Appl
8	40	75.5	1893	US-08-188-582-11	Sequence 11, Appl
9	40	75.5	1893	US-08-646-715-11	Sequence 11, Appl
10	38.5	72.6	115	US-09-461-325-160	Sequence 160, App
11	38.5	72.6	115	US-09-996-243-95	Sequence 95, Appl
12	37	69.8	350	US-08-495-695B-33	Sequence 33, Appl
13	37	69.8	375	US-08-175-412-2	Sequence 2, Appli
14	37	69.8	375	US-08-555-268A-2	Sequence 2, Appli
15	37	69.8	375	US-08-495-695B-2	Sequence 2, Appli
16	37	69.8	375	US-09-200-673-17	Sequence 17, Appl
17	37	69.8	375	PCT-US94-14436-2	Sequence 2, Appli
18	37	69.8	470	US-09-252-991A-28702	Sequence 28702, A
19	36	67.9	16	US-09-177-249-113	Sequence 113, App
20	36	67.9	287	US-09-690-454-55	Sequence 55, Appl
21	36	67.9	417	US-09-252-991A-20165	Sequence 20165, A
22	36	67.9	1209	US-09-252-991A-25844	Sequence 25844, A
23	35.5	67.0	106	US-09-634-238-349	Sequence 349, App
24	35	66.0	128	US-08-905-223-306	Sequence 306, App
25	35	66.0	200	US-09-252-991A-31738	Sequence 31738, A
26	35	66.0	219	US-09-252-991A-32844	Sequence 32844, A
27	35	66.0	241	US-08-634-475-7	Sequence 7, Appli

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
PCT-US95-04682-4
Query Match 75.5%; Score 40; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHH 5
DB 4 WPFHH 8
RESULT 4
US-09-461-325-465
; Sequence 465, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461.325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 465
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-465
Query Match 75.5%; Score 40; DB 4; Length 128;
Best Local Similarity 71.4%; Pred. No. 8.1;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WPFHH 7
DB 50 WPFHH 56
RESULT 5
US-09-461-325-464
; Sequence 464, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461.325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-227-536-4
Query Match 75.5%; Score 40; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHH 5
DB 4 WPFHH 8
RESULT 3
PCT-US95-04682-4
; Sequence 4, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308XG999
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids

/ EARLIER FILING DATE: 1998-06-16
 / EARLIER APPLICATION NUMBER: 60/089,509
 / EARLIER FILING DATE: 1998-06-16
 / EARLIER APPLICATION NUMBER: 60/089,510
 / EARLIER FILING DATE: 1998-06-16
 / EARLIER APPLICATION NUMBER: 60/090,112
 / EARLIER FILING DATE: 1998-06-22
 / EARLIER APPLICATION NUMBER: 60/090,113
 / EARLIER FILING DATE: 1998-06-22
 / NUMBER OF SEQ ID NOS: 532
 / SOFTWARE: Patentin Ver. 2.0
 / SEQ ID NO 464
 / LENGTH: 165
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-461-325-464

Query Match 75.5%; Score 40; DB 4; Length 165;
 Best Local Similarity 71.4%; Pred. No. 10;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
 |||||
 Db 87 WPHRRH 93

RESULT 6
 US-08-188-582-14
 / Sequence 14, Application US/08188582
 / Patent No. 5534410
 / GENERAL INFORMATION:
 / APPLICANT: Tjian, Robert
 / APPLICANT: Comai, Lucio
 / APPLICANT: Dynlacht, Brian D.
 / APPLICANT: Hoey, Timothy
 / APPLICANT: Ruppert, Siegfried
 / APPLICANT: Tanese, Naoko
 / APPLICANT: Wang, Edith
 / APPLICANT: Weinzierl, Robert O.J.
 / TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 / NUMBER OF SEQUENCES: 36
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 / STREET: 4 Embarcadero Center, Suite 3400
 / CITY: San Francisco
 / STATE: California
 / COUNTRY: USA
 / ZIP: 94111-4187
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.25
 / CURRENT APPLICATION NUMBER: US/08/188,582
 / FILING DATE: 28-JAN-1994
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Osman, Richard A
 / REGISTRATION NUMBER: 36,627
 / REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
 / TELEPHONE: (415) 781-1989
 / TELEFAX: (415) 398-3249
 / INFORMATION FOR SEQ ID NO: 14:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1872 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: peptide

US-08-188-582-14
 Query Match 75.5%; Score 40; DB 1; Length 1872;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
 |||||
 Db 1526 WPFHH 1530

RESULT 7
 US-08-646-715-14
 / Sequence 14, Application US/08646715
 / Patent No. 5637666
 / GENERAL INFORMATION:
 / APPLICANT: Tjian, Robert
 / APPLICANT: Comai, Lucio
 / APPLICANT: Dynlacht, Brian D.
 / APPLICANT: Hoey, Timothy
 / APPLICANT: Ruppert, Siegfried
 / APPLICANT: Tanese, Naoko
 / APPLICANT: Wang, Edith
 / APPLICANT: Weinzierl, Robert O.J.
 / TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 / NUMBER OF SEQUENCES: 36
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 / STREET: 4 Embarcadero Center, Suite 3400
 / CITY: San Francisco
 / STATE: California
 / COUNTRY: USA
 / ZIP: 94111-4187
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.25
 / CURRENT APPLICATION NUMBER: US/08/646,715
 / FILING DATE: 09-MAY-1996
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/188,582
 / FILING DATE: 28-JAN-1994
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Osman, Richard A
 / REGISTRATION NUMBER: 36,627
 / REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
 / TELEPHONE: (415) 781-1989
 / TELEFAX: (415) 398-3249
 / INFORMATION FOR SEQ ID NO: 14:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1872 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: peptide

Query Match 75.5%; Score 40; DB 1; Length 1872;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
 |||||
 Db 1526 WPFHH 1530

RESULT 8

US-08-188-582-11
; Sequence 11, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1893 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-715-11
Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHH 5
DB 1547 WPFHH 1551
RESULT 10
US-09-461-325-160
; Sequence 160, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461.325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 160
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens

US-08-188-582-11
; Sequence 11, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1893 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-582-11
Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHH 5
DB 1547 WPFHH 1551
RESULT 9
US-08-646-715-11
; Sequence 11, Application US/08646715
; Patent No. 5637686
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:

```
;
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-461-325-160

Query Match          72.8%; Score 38.5; DB 4; Length 115;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 WP-----FHHR 7
Db 59 WPFRRGHLCIFHHR 74

RESULT 11
US-09-996-243-95
; Sequence 95, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavins, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Paul, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996.243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
```

; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07

; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09
 Query Match 72.6%; Score 38.5; DB 4; Length 115;
 Best Local Similarity 43.8%; Pred. No. 13; Indels 9; Gaps 1;
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;
 QY 1 WP-----FHHHR 7
 DB 59 WPFRRRCHLCIFHHHR 74
 RESULT 12
 US-08-495-695B-33
 ; Sequence 33, Application US/08495695B
 ; Patent No. 5976814
 ; GENERAL INFORMATION:
 ; APPLICANT: Bard, Jonathan A.
 ; APPLICANT: Walker, Mary
 ; APPLICANT: Branchek, Theresa
 ; APPLICANT: Weinschenk, Richard L.
 ; TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
 ; TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
 ; TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/495.695B
 ; FILING DATE: 13-Jan-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 44743-A-PCT-US/JPW/JHB
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; INFORMATION FOR SEQ ID NO: 33:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 350 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-495-695B-33
 Query Match 69.8%; Score 37; DB 2; Length 350;
 Best Local Similarity 71.4%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WPFHHHR 7
 DB 205 WPLAHR 211
 RESULT 13
 US-08-176-412-2
 ; Sequence 2, Application US/08176412
 ; Patent No. 5516653
 ; GENERAL INFORMATION:
 ; APPLICANT: Bard, Jonathan A.
 ; APPLICANT: Walker, Mary
 ; APPLICANT: Branchek, Theresa

APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,412
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743/JPW/TEP
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-412-2

Query Match 69.8%; Score 37; DB 1; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
DB 205 WPLAHR 211

RESULT 14
US-08-555-268A-2
Sequence 2, Application US/08555268A
Patent No. 5958709
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
APPLICANT: Walker, Mary
APPLICANT: Branchek, Theresa
APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,268A

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743-Z/JPW/MAT
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-555-268A-2

Query Match 69.8%; Score 37; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
DB 205 WPLAHR 211

RESULT 15
US-08-495-695B-2
Sequence 2, Application US/08495695B
Patent No. 5976814
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
APPLICANT: Walker, Mary
APPLICANT: Branchek, Theresa
APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/495,695B
FILING DATE: 13-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743-A-PCT-US/JPW/JHB
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-495-695B-2

Query Match 69.8%; Score 37; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 WPFHHR 7
Db 205 WPLAHR 211

Search completed: February 11, 2004, 17:13:38
Job time : 11.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	7	23	AAE14554 Human alpha-synuc
2	42	79.2	151	21	AAG36521 Arabidopsis thalia
3	42	79.2	176	21	AAG20314 Arabidopsis thalia
4	42	79.2	176	21	AAG38740 Arabidopsis thalia
5	42	79.2	181	21	AAG20313 Arabidopsis thalia
6	42	79.2	181	21	AAG38739 Arabidopsis thalia
7	42	79.2	309	22	AAB87830 Amino acid sequenc
8	42	79.2	310	21	AAG20312 Arabidopsis thalia
9	42	79.2	310	21	AAG38738 Arabidopsis thalia

10	42	79.2	541	22	AB568621 Drosophila melanog
11	40	75.5	123	22	AA004423 Human polypeptide
12	40	75.5	128	21	AA86527 Human gene 72-enco
13	40	75.5	165	21	AA86526 Human gene 72-enco
14	40	75.5	898	23	ABG70103 Human prey protein
15	40	75.5	1575	23	AAU84357 Protein TAF2A diff
16	40	75.5	1872	15	AAU86493 TATA-binding prote
17	40	75.5	1872	17	AAW06078 Drosophila TATA-bi
18	40	75.5	1872	18	AAW25030 TATA-binding prote
19	40	75.5	1886	22	ABG06009 Novel human diago
20	40	75.5	1893	17	AAW56491 Human TATA-binding
21	40	75.5	1893	15	AAW06082 TATA-binding prote
22	40	75.5	1893	18	AAW25020 Novel human diago
23	40	75.5	1924	22	ABG06008 Novel human diago
24	39	73.6	11	23	AAU97110 pAMG21-human MK61
25	38.5	72.6	115	21	AAU86243 Human secreted pro
26	38.5	72.6	115	21	AAU86649 Membrane-bound pro
27	38.5	72.6	115	22	AAU29066 Human PRO polypept
28	38.5	72.6	115	22	AAU29066 Human PRO537 (UNQ3
29	38.5	72.6	115	24	ABU71154 Human PRO537 prote
30	38.5	72.6	115	24	ABU65611 Human secreted/tra
31	38.5	72.6	115	24	ABU65944 Novel human secret
32	38.5	72.6	115	24	ABU87448 Human secreted/tra
33	38.5	72.6	115	24	ABU85306 Human PRO polypept
34	38.5	72.6	115	24	ABU59085 Novel human secret
35	38.5	72.6	115	24	ABU59212 Novel human secret
36	38.5	72.6	115	24	ABU59361 Human secreted/tra
37	38.5	72.6	115	24	ABU60496 Human secreted/tra
38	38.5	72.6	115	24	ABU57987 Human PRO polypept
39	38.5	72.6	115	24	ABU58442 Human PRO polypept
40	38.5	72.6	115	24	ABU58918 Human secreted/tra
41	38.5	72.6	115	24	ABU5978 Human secreted/tra
42	38.5	72.6	115	24	ABU56973 Human PRO polypept
43	38.5	72.6	115	24	ABU13878 Human PRO537 polyp
44	38.5	72.6	115	24	ABU10552 Human secreted/tra
45	38.5	72.6	115	24	ABU10833 Human PRO polypept

ALIGNMENTS

RESULT 1

AAE14554
ID AAE14554 standard; peptide; 7 AA.

XX AAE14554;

AC AAE14554;

DT 17-MAY-2002 (first entry)

XX Human alpha-synuclein aggregation inhibitor #9.
XX Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
XX Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
XX multiple system atrophy; Hallervorden-Spatz disease; human.

OS Homo sapiens

PN WO200204482-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21379.

XX 07-JUL-2000; 2000US-217319P.

XX 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

XX Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

XX Determination of an agent capable of inhibiting aggregation of alpha

PT

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
XX Claim 40; Page 37; 52pp; English.
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
XX Sequence 7 AA;

Query Match 100.0%; Score 53; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WPFHHR 7
| | | | |
Db 1 WPFHHR 7

RESULT 2
AAG36521
ID AAG36521 standard; Protein; 151 AA.
XX
AC AAG36521;
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44768.

XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144632.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.
99US-0145087.
99US-0145089.
99US-0145192.
99US-0145145.
99US-0145218.
99US-0145224.

PR 14-MAY-1999;
PR 14-MAY-1999;
PR 14-MAY-1999;
PR 19-MAY-1999;
PR 20-MAY-1999;
PR 21-MAY-1999;
PR 24-MAY-1999;
PR 25-MAY-1999;
PR 27-MAY-1999;
PR 28-MAY-1999;
PR 01-JUN-1999;
PR 03-JUN-1999;
PR 04-JUN-1999;
PR 07-JUN-1999;
PR 08-JUN-1999;
PR 10-JUN-1999;
PR 10-JUN-1999;
PR 14-JUN-1999;
PR 16-JUN-1999;
PR 16-JUN-1999;
PR 17-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 21-JUN-1999;
PR 22-JUN-1999;
PR 23-JUN-1999;
PR 23-JUN-1999;
PR 24-JUN-1999;
PR 28-JUN-1999;
PR 29-JUN-1999;
PR 30-JUN-1999;
PR 01-JUL-1999;
PR 01-JUL-1999;
PR 02-JUL-1999;
PR 06-JUL-1999;
PR 08-JUL-1999;
PR 09-JUL-1999;
PR 12-JUL-1999;
PR 13-JUL-1999;
PR 14-JUL-1999;
PR 15-JUL-1999;
PR 16-JUL-1999;
PR 16-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 20-JUL-1999;
PR 20-JUL-1999;
PR 21-JUL-1999;
PR 21-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147433.
PR 08-AUG-1999; 99US-0147435.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147204.
PR 10-AUG-1999; 99US-0147302.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157723.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 151;
Best Local Similarity 83.3%; Pred. No. 8.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 4 WPLHHH 9

RESULT 3
AAG20314
ID AAG20314 standard; Protein; 176 AA.
XX
AC AAG20314;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22452.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
```


PR 28-OCT-1999; 99US-01611992.
PR 28-OCT-1999; 99US-01611993.
PR 29-OCT-1999; 99US-0162142.
Query March 79.2%; Score 42; DB 21; Length 176;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WPFHHH 6
DB 29 WPLHHH 34
RESULT 4
AAG38740
ID AAG38740 standard; Protein; 176 AA.
XX
AC AAG38740;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47837.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
KW
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

Query Match 79.2%; Score 42; DB 21; Length 176;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	21-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140655.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144632.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149923.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0158293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 79.2%; Score 42; DB 21; Length 181;		
Best Local Similarity 83.3%; Pred. No. 11;		
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1 WPFHHH 6	
Db	34 WPLHHH 39	
RESULT 6		

[illegible]

```

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 181;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 34 WPLHHH 39

RESULT 7
AAB67830
ID AAB67830 standard; Protein; 309 AA.
XX
AC AAB67830;
XX
DT 29-JUN-2001 (first entry)
XX

```

```

DE Amino acid sequence of a plant transcription factor G234.
XX Transcription factor; flowering time; transgenic plant; vernalisation;
KW plant development; plant physiology; flowering.
XX Arabidopsis thaliana.
XX WO200126459-A2.
XX 19-APR-2001.
XX 11-OCT-2000; 2000WO-US28141.
XX 12-OCT-1999; 99US-0159464.
XX 08-NOV-1999; 99US-0164132.
XX 17-NOV-1999; 99US-0166228.
XX 17-APR-2000; 2000US-0197899.
XX 22-AUG-2000; 2000US-0227439.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (RATC/) RATCLIFFE O.
PA (HEAR/) HEARD J.
PA (SAMA/) SAMARA R.
PA (CREE/) CREELMAN R.
PA (KEDD/) KEDDIE J.
PA (JIAN/) JIANG C.
PA (REUB/) REUBER L.
PA (RIEC/) RIECHMANN J L.
XX
PI Ratcliffe O, Heard J, Samaha R, Creelman R, Keddie J, Jiang C;
PI Reuber L, Riechmann JL;
XX WPI; 2001-266398/27.
XX N-PSDB; AAF80400.
XX
PT New transgenic plant comprises a recombinant polynucleotide encoding a
PT plant transcription factor polypeptide and has a modified flowering
PT time or vernalization requirement -
XX
XX Example 7; Page 73-75; 109pp; English.
XX
XX The present sequence represents a plant transcription factor protein
XX which modifies the flowering time of a plant. The polynucleotide
XX sequence is used to produce transgenic plants which have a modified
XX flowering time or a modified vernalisation requirement. The
XX polynucleotides and polypeptides are useful for modifying plant
XX development, physiology or biochemistry such that the modified plants
XX have a trait advantage over wild type plants. In particular they are
XX useful for accelerating, delaying or preventing flowering. The
XX polynucleotides are also useful as nucleic acid probes and primers.
XX They may be used to identify proteins that can modify the activity of
XX the transcription factor.
XX
XX Sequence 309 AA;

Query Match 79.2%; Score 42; DB 22; Length 309;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 8
AAG20312
ID AAG20312 standard; Protein; 310 AA.
XX
AC AAG20312;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22450.

```

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 03-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134768.

XX 20-MAY-1999; 99US-0134941.

XX 21-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144336.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.

PR	31-AUG-1999;	99US-0151438.	PN	EPI033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070.	XX	
PR	13-SEP-1999;	99US-0153258.	PP	
PR	15-SEP-1999;	99US-0154018.	XX	
PR	16-SEP-1999;	99US-0154039.	XX	25-FEB-2000; 2000EP-0301439.
PR	20-SEP-1999;	99US-0154779.	PR	25-FEB-1999;
PR	22-SEP-1999;	99US-0155139.	PR	05-MAR-1999;
PR	23-SEP-1999;	99US-0155486.	PR	09-MAR-1999;
PR	24-SEP-1999;	99US-0155659.	PR	23-MAR-1999;
PR	28-SEP-1999;	99US-0156458.	PR	25-MAR-1999;
PR	29-SEP-1999;	99US-0156596.	PR	29-MAR-1999;
PR	04-OCT-1999;	99US-0157117.	PR	01-APR-1999;
PR	05-OCT-1999;	99US-0157753.	PR	06-APR-1999;
PR	06-OCT-1999;	99US-0157865.	PR	08-APR-1999;
PR	07-OCT-1999;	99US-0158029.	PR	16-APR-1999;
PR	08-OCT-1999;	99US-0158232.	PR	19-APR-1999;
PR	12-OCT-1999;	99US-0158369.	PR	21-APR-1999;
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999;
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999;
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999;
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999;
PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999;
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999;
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999;
PR	18-OCT-1999;	99US-0159838.	PR	07-MAY-1999;
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;
PR	25-OCT-1999;	99US-0161404.	PR	24-MAY-1999;
PR	25-OCT-1999;	99US-0161405.	PR	25-MAY-1999;
PR	25-OCT-1999;	99US-0161406.	PR	27-MAY-1999;
PR	26-OCT-1999;	99US-0161359.	PR	28-MAY-1999;
PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999;
PR	26-OCT-1999;	99US-0161361.	PR	03-JUN-1999;
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999;
PR	28-OCT-1999;	99US-0161902.	PR	07-JUN-1999;
PR	28-OCT-1999;	99US-0161903.	PR	08-JUN-1999;
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;
Query Match 79.2%; Score 42; DB 21; Length 310;			PR	14-JUN-1999;
Best Local Similarity 83.3%; Pred. No. 19;			PR	16-JUN-1999;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			PR	16-JUN-1999;
OY	1 WPFHHH 6		PR	17-JUN-1999;
DB	163 WFLHHH 168		PR	18-JUN-1999;
RESULT 9			PR	18-JUN-1999;
ID	AAAG38738		PR	18-JUN-1999;
XX	AAAG38738 standard; Protein; 310 AA.		PR	18-JUN-1999;
AC	AAAG38738;		PR	18-JUN-1999;
XX	DT		PR	18-JUN-1999;
XX	DT 18-OCT-2000 (first entry)		PR	18-JUN-1999;
XX	AAAG38738;		PR	18-JUN-1999;
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 47835.		PR	18-JUN-1999;
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 47835.		PR	18-JUN-1999;
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	18-JUN-1999;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	18-JUN-1999;
XX	termination sequence.		PR	18-JUN-1999;
OS	Arabidopsis thaliana.		PR	18-JUN-1999;
XX			PR	18-JUN-1999;
XX			PR	18-JUN-1999;

```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 05-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149732.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 310;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 10
ABB68621
ID ABB68621 standard; Protein; 541 AA.
XX
AC ABB68621;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32655.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
```

```

XX PA (PEKE ) PE CORP NY.
XX PI
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL112724.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX PS Disclosure; SEQ ID NO 32655; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 541 AA;
XX
XX Query Match 79.2%; Score 42; DB 22; Length 541;
XX Best Local Similarity 83.3%; Pred. No. 33;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 WPFHH 6
XX |||||
XX DB 300 WPLHH 305
XX
XX RESULT 11
XX AAO04423
XX ID AAO04423 standard; Protein; 123 AA.
XX AC AAO04423;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 18315.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX OS Homo sapiens.
XX
XX PN WO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX PF 26-FEB-2001; 2001WO-US04927.
XX
XX PR 28-FEB-2000; 2000US-0515126.
XX
XX PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX DR N-PSDB; AA184334.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
XX Claim 20; SEQ ID NO 18315; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 123 AA;
XX
XX Query Match 75.5%; Score 40; DB 22; Length 123;
XX Best Local Similarity 83.3%; Pred. No. 16;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 WPFHH 6
XX |||||
XX DB 57 WPFHH 62
XX
XX RESULT 12
XX AAY86527
XX ID AAY86527 standard; Protein; 128 AA.
XX AC AAY86527;
XX DT 19-APR-2000 (first entry)
XX DE Human gene 72-encoded protein fragment, SEQ ID NO:442.
XX
XX Human; secreted protein; cancer; tumour; developmental abnormality;
XX foetal deficiency; blood disorder; immune system disorder; inflammation;
XX autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
XX therapy.
XX
XX OS Homo sapiens.
XX
XX PN WO9966041-A1.
XX
XX PD 23-DEC-1999.
XX
XX PF 15-JUN-1999; 99WO-US13418.
XX
XX PR 16-JUN-1998; 98US-0089507.
XX
XX PR 16-JUN-1998; 98US-0089508.
XX
XX PR 16-JUN-1998; 98US-0089509.
XX
XX PR 16-JUN-1998; 98US-0089510.
XX
XX PR 22-JUN-1998; 98US-0090112.
XX
XX PR 22-JUN-1998; 98US-0090113.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;
XX Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R;
XX Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G;
XX
XX WPI; 2000-106100/09.
XX
XX New isolated human genes and the secreted polypeptides they encode,
XX PT useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX

```

DR WPI: 2000-106100/09.

XX New isolated human genes and the secreted polypeptides they encode,

PT useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders -

XX

XX Disclosure; Page 152; 586pp; English.

XX

XX AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.

CC AAZ96215 to AAY86333 are the secreted proteins encoded by the 94 human

CC genes. This sequence represents a fragment of one of the human secreted

CC proteins. The genes and their corresponding secreted polypeptides are

CC useful for preventing, treating or ameliorating medical conditions,

CC e.g., by protein or gene therapy. Also pathological conditions can be

CC diagnosed by determining the amount of the new polypeptides in a sample

CC or by determining the presence of mutations in the new genes. Specific

CC uses are described for each of the 94 genes, based on which tissues they

CC are most highly expressed in, and include developing products for the

CC diagnosis or treatment of cancer, tumours, developmental abnormalities

CC and foetal deficiencies, blood disorders, diseases of the immune system,

CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive

CC disorders, schizoprenia, arthritis, asthma, psoriasis, sepsis, skin

CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney

CC disorders, digestive/endocrine disorders, infections and AIDS. The

CC polypeptides are also useful for identifying their binding partners.

CC The sequences shown in AAY86334 to AAY86585 represent fragments of the

CC secreted proteins.

XX

XX Sequence 165 AA;

XX

XX Query Match 75.5%; Score 40; DB 21; Length 165;

XX Best Local Similarity 71.4%; Pred. No. 21;

XX Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0

QY 1 WPFPHHR 7
|||:|

DB 87 WPNRHR 93

RESULT 14

ABG70103

ID ABG70103 standard; Protein; 898 AA.

XX ABG70103;

XX

XX 21-OCT-2002 (first entry)

XX

XX Human prey protein for Shigella ospC1 #21.

XX

XX Prey protein; ospB; ospD1; ipaB; ipaC; ipaH9.8; ospG;

KW ospC1; Shigella; shigellosis; bacillary dysentery; antibacterial;

KW yeast two-hybrid system; protein-protein interaction; SID;

KW selected interacting domain; human.

XX

XX Homo sapiens.

XX

XX WO200257303-A2.

XX

XX 25-JUL-2002.

XX

XX 11-JAN-2002; 2002WO-BP00777.

XX

XX 12-JAN-2001; 2001US-261130P.

PR

XX (HYBR-) HYBRIGENICS.

XX

XX Legrain P;

XX

XX WPI: 2002-599706/64.

DR

DR N-PSDB; ABS51496.

XX

XX New complex of protein-protein interactions between a bait Shigella

PT flexneri polypeptide and a prey mammalian or human placental polypeptide

Search completed: February 11, 2004, 17:03:02
Job time : 42.25 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-10

Perfect score: 45

Sequence: 1 HLYHHT 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SID1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SID1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	7	23	Human alpha-synuclein
2	39	86.7	98	21	Arabidopsis thaliana
3	37	82.2	83	22	Propionibacterium
4	37	82.2	140	22	Drosophila melanogaster
5	37	82.2	442	22	Drosophila melanogaster
6	36	80.0	91	23	Lactococcus lactis
7	36	80.0	110	17	Tomato P19 protein
8	36	80.0	136	24	Lolium perenne LpA
9	36	80.0	136	24	Lolium perenne ASR

10	36	80.0	140	23	AAW52548	Maize ASR protein.
11	36	80.0	210	23	AAU87835	T. aureum 7091 par
12	36	80.0	272	23	AAU87836	T. aureum 7091 elo
13	36	80.0	272	23	AAU87837	T. aureum 7091 elo
14	36	80.0	272	23	AAU87839	T. aureum 7091 elo
15	36	80.0	272	23	AAU87841	T. aureum 7091 elo
16	36	80.0	272	23	AAU87842	T. aureum 7091 elo
17	36	80.0	283	21	AAV79250	Mammalian putative
18	36	80.0	285	9	AAV82590	polypeptide with g
19	36	80.0	299	23	AAU87834	Mouse elongase MEL
20	36	80.0	750	24	ABU07373	Human protein NOV5
21	36	80.0	873	24	ABU11589	Human MDDT polyep
22	36	80.0	1036	24	ABU11770	Human MDDT polyep
23	35	77.8	59	21	AAU41234	Human ORF3289 prot
24	35	77.8	76	23	ABP34316	Novel human diagno
25	35	77.8	113	22	ABG11726	Human ORF3899 prot
26	35	77.8	119	23	ABP34926	Arabidopsis thaliana
27	35	77.8	176	21	AAU42058	Arabidopsis thaliana
28	35	77.8	187	21	AAU42057	Arabidopsis thaliana
29	35	77.8	206	23	ABG79615	Sunflower ribonuc
30	35	77.8	242	11	AAU06483	18-2-3-IR159. Mu
31	35	77.8	277	22	ABG61428	Drosophila melanog
32	35	77.8	292	21	AAU42056	Arabidopsis thaliana
33	35	77.8	294	22	ABG24169	Novel human diagno
34	35	77.8	311	22	AAU65563	Human carboxypepti
35	35	77.8	349	17	AAU06182	Modified HCPB (D25
36	35	77.8	349	17	AAU06173	mature HCPB (H18)6
37	35	77.8	349	17	AAU06181	Modified HCPB (D25
38	35	77.8	349	18	AAU13750	Carboxypeptidase B
39	35	77.8	349	18	AAU13751	Carboxypeptidase B
40	35	77.8	349	18	AAU13752	Carboxypeptidase B
41	35	77.8	349	18	AAU13753	Carboxypeptidase B
42	35	77.8	349	18	AAU13757	Carboxypeptidase B
43	35	77.8	349	18	AAU13758	Carboxypeptidase B
44	35	77.8	349	18	AAU13759	Carboxypeptidase B
45	35	77.8	349	18	AAU13760	Carboxypeptidase B

ALIGNMENTS

RESULT 1

AAE14555
ID AAE14555 standard; peptide; 7 AA.

XX AAE14555;

XX 17-MAY-2002 (first entry)

XX Human alpha-synuclein aggregation inhibitor #10.

XX Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
XX Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
XX multiple system atrophy; Hallervorden-Spatz disease; human.

XX Homo sapiens.

XX WO200204482-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21379.

XX 07-JUL-2000; 2000US-217319P.

XX 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

XX Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

XX Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
XX Claim 40; Page 37; 52pp; English.
XX
XX The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
XX Sequence 7 AA;

Query Match 100.0%; Score 45; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HLYHHKT 7
|||||
Db 1 HLYHHKT 7

RESULT 2
AAG56024
ID AAG56024 standard; Protein; 98 AA.

XX AAG56024;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 71943.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136352.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 23-JUL-1999; 99US-0145224.

```

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match      86.7%; Score 39; DB 21; Length 98;
Best Local Similarity 71.4%; Pred. No. 6.5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHKT 7
DB 84 HIYHQT 90

RESULT 3
AAU57344
ID AAU57344 standard; Protein; 83 AA.
XX
AC AAU57344;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #18240.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59582.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 18539; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies

```

specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and to treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct/sequences.

Query Match 82.2%; Score 37; DB 22; Length 83;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	HL	YHK	6
			:	
Dy	49	HL	YHR	54

RESULT 4
ABB65911
ID ABB65911 standard: Protein: 140 AA.

AC ABB65911;

XX
DT 26-MAR-2002 (first entry)

DE
XX
XX

XX
Drosophila; developmental biology; cell signalling; insecticide;
KW
pharmaceutical.

OS *Drosophila melanogaster*.

XX
PN
WQ200171042-A2.

XX 27 FEB 2007

XX
PF 23-MAR-2001: 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

PK 23-MAR-2000; 2000US-151037E;
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI
Venter JC, Adams M, Li PWD, Myers EW;

XX
DB WPT: 2001-656860/75.

DR N-PSDB; ABL10014.

AA New isolated nucleic acid detection reagent for detecting 1000 or more
PT PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT PT interactions -
PT PT

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from proscophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

AX	Sequence	140 AA;
SO		

Query Match	82.2%;	Score 37;	DB 22;	Length 140;
Best Local Similarity	71.4%;	Pred. No. 22;		

	Matches	5; Conservative	1; Mismatches	1; Indels	0; Gaps
QY	1	HLVHKT	7		
		: :			
DB	37	HVYHST	43		

RESULT 5
ABB62210
IN ABB62210 standard: protein: 442 AA.

AC ABB62210:

26-MAR-2002 (first entry)

XX
Ducanville polycarbonate SEC ID NO 13422

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical

8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

XX
XX
.

4
4
5
4
4
5
5
4
5
=

4 X
4 X

FD 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US09231.
YY

23-MAR-2000; 2000US-191637P.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

EA	(LAW) / IS CORE MIA		
XX			
BT	Wentworth TO Adams M	Y; DWT	Mvaya EW

XX
PI
venter JC, Adams M,

DR WPI; 2001-656860/

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
PT

XX Disclosure: SEQ ID NO 13422; 21bp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA

CC sequences (ABU01840-ABU16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at <http://www.int/pub/published/seq>.

CC

XX Sequence 442 AA:

Query Match	82.2%	Score 37;	DB 22;	Length 442;
Best Local Similarity	71.4%;	Pred. No. 78;		
Matches 5;	Conservative	1;	Mismatches	1;
			Indels	0;
			Gaps	0;

Qy	1	HLVHKT	7
		: :	
Db	320	HVYHHST	326

RESULT 6
ABB53926

XX

AC ABB53926
XX

DT 16-MAY-2002 (first entry)
 XX Lactococcus lactis protein ygaE.
 DE
 XX
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 KW
 XX Lactococcus lactis IL1403.
 OS
 XX FR2807446-A1.
 FN
 XX 12-OCT-2001.
 PD
 XX
 XX 11-APR-2000; 2000FR-0004630.
 PF
 XX 11-APR-2000; 2000FR-0004630.
 PR
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 PI
 XX WPI; 2002-043418/06.
 DR
 XX New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species -
 FT
 XX Claim 6; SEQ ID No 628; 2504pp; French.
 PS
 XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (AB90521) and related proteins (ABBS3300-ABBS5621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 91 AA;
 Query Match 80.0%; Score 36; DB 23; Length 91;
 Best Local Similarity 85.7%; Pred.No. 21;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HLYHHKT 7
 DB 56 HLCCHKT 62
 |||||
 |||||
 RESULT 7
 AAW02631
 ID AAW02631 standard; Protein; 110 AA.
 XX
 XX AAW02631;
 AC
 XX 07-NOV-1996 (first entry)
 DT
 XX Tomato P119 protein.
 DE
 XX P119 gene; promoter; fruit-specific expression; transgenic plant;
 KW herbicide resistance; disease resistance; crop protection; tomato;
 KW tobacco.
 XX
 XX Lycopersicon esculentum cv. Golden Nugget.
 OS
 XX WO9619103-A1.
 FN
 XX 27-JUN-1996.
 PD
 XX 29-NOV-1995; 95WO-US15482.
 PF
 XX 20-DEC-1994; 94US-0359696.
 PR

XX (DNAP) DNA PLANT TECHNOLOGY CORP.
 PA
 XX Dunsmuir P, Stott JS;
 PI
 XX WPI; 1996-309200/31.
 DR
 XX N-PSDB; AAT32863.
 DR
 XX New plant P119 promoter - useful for generating transgenic plant,
 PT pref. tomato or tobacco, with e.g. herbicide, fungal disease or
 PT bacterial disease resistance
 PT
 XX Claim 17; Page 41; 50pp; English.
 PS
 XX P119 (AAW02631) is the product of a novel gene (see also AAT32863)
 CC isolated from the pericarp of cherry tomato cv. Golden Nugget.
 CC P119 mRNA is present at high levels in tomato pericarp (both
 CC mature and at various stages of ripening), at moderate levels in
 CC flowers and immature green fruit, at low levels in roots and stems,
 CC and is not present in leaves. The promoter (see also AAT32864) of
 CC the P119 gene can be used to express operably linked DNA
 CC sequences primarily in fruit tissue of transgenic plants, esp.
 CC tomato and tobacco.
 CC
 XX Sequence 110 AA;
 SQ
 Query Match 80.0%; Score 36; DB 17; Length 110;
 Best Local Similarity 83.3%; Pred.No. 26;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HLYHHK 6
 DB 10 HLFHHK 15
 |||||
 |||||
 RESULT 8
 AAE34013
 ID AAE34013 standard; Protein; 136 AA.
 XX
 XX AAE34013;
 AC
 XX 02-MAY-2003 (first entry)
 DT
 XX Lolium perenne LpASR protein.
 DE
 XX Abscisic acid-inducible and stress responsive protein; ASR; A22; PKABA;
 KW stress-inducible cysteine protease; late embryogenesis abundant protein;
 KW LEA; dehydrin; DHN; abscisic acid-induced protein kinase; gene therapy;
 KW CYS; seed development; plant tolerance; germination; plant protectant;
 KW ryegrass.
 XX
 XX Lolium perenne.
 OS
 XX WO200290547-A1.
 FN
 XX 14-NOV-2002.
 PD
 XX 07-MAY-2002; 2002WO-AU00564.
 PF
 XX 07-MAY-2001; 2001AU-0004821.
 PR
 XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
 PA (AGRE-) AGRESEARCH LTD.
 PA
 XX Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
 PI
 XX WPI; 2003-129183/12.
 DR
 XX N-PSDB; AAD52414.
 DR
 XX New isolated nucleic acid encoding ASR, A22, CYS, LEA, DHN or PKABA
 PT proteins, useful as molecular genetic markers, and in modifying plant
 PT and/or seed development and responses to stresses and adverse
 PT environmental stimuli -

XX Claim 24; Fig 2; 231pp; English.

XX The invention relates to nucleic acid encoding abscisic acid-inducible

CC and stress responsive proteins (ASR and A22), stress-inducible cysteine

CC proteases (CYS), late embryogenesis abundant proteins (LEA), dehydrins

CC (DHN) and abscisic acid-induced protein kinases (PKABA). The invention

CC also relates to a method for modification of plant and seed development

CC as molecular genetic markers. The method is useful for modifying plant

CC response to an environmental stimulus, modifying plant tolerance to

CC abiotic, osmotic and/or temperature stresses, modifying seed dormancy

CC and/or germination, development, maturation, and modifying a plant

CC developmental process. They are also useful for modifying plant

CC tolerance and adaptation to stresses and adverse environmental stimuli.

CC The invention is also used in gene therapy. The present sequence is

CC Lolium perenne LpASR protein.

XX Sequence 136 AA;

SQ Query Match 80.0%; Score 36; DB 24; Length 136;

Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

Db 8 HLFHHK 13

||:||||

RESULT 9

AAE34023

ID AAE34023 standard; Protein; 136 AA.

AC AAE34023;

XX 02-MAY-2003 (first entry)

DT Lolium perenne ASRa protein.

DE Abscisic acid-inducible and stress responsive protein; ASR; A22; PKABA;

XX stress-inducible cysteine protease; late embryogenesis abundant protein;

XX LEA; dehydrin; DHN; abscisic acid-induced protein kinase; gene therapy;

XX CYS; seed development; plant tolerance; germination; plant protectant;

XX ryegrass.

XX Lolium perenne.

OS WO200290547-A1.

PN 14-NOV-2002.

XX 07-MAY-2002; 2002WO-AU00564.

XX 07-MAY-2001; 2001AU-0004821.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

PA (AGRE-) AGRESEARCH LTD.

XX Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;

PI WPI; 2003-129183/12.

DR N-PSDB; AAD52469.

XX New isolated nucleic acid encoding ASR, A22, CYS, LEA, DHN or PKABA

PT proteins, useful as molecular genetic markers, and in modifying plant

PT and/or seed development and responses to stresses and adverse

PT environmental stimuli -

XX Claim 24; Fig 35; 231pp; English.

XX The invention relates to nucleic acid encoding abscisic acid-inducible

CC and stress responsive proteins (ASR and A22), stress-inducible cysteine

CC proteases (CYS), late embryogenesis abundant proteins (LEA), dehydrins

CC (DHN) and abscisic acid-induced protein kinases (PKABA). The invention

CC also relates to a method for modification of plant and seed development

CC as molecular genetic markers. The method is useful for modifying plant

CC response to an environmental stimulus, modifying plant tolerance to

CC abiotic, osmotic and/or temperature stresses, modifying seed dormancy

CC and/or germination, development, maturation, and modifying a plant

CC developmental process. They are also useful for modifying plant

CC tolerance and adaptation to stresses and adverse environmental stimuli.

CC The invention is also used in gene therapy. The present sequence is

CC Lolium perenne LpASR protein.

XX Sequence 136 AA;

SQ Query Match 80.0%; Score 36; DB 24; Length 136;

Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

Db 8 HLFHHK 13

||:||||

RESULT 10

AAE34023

ID AAE34023 standard; Protein; 140 AA.

AC AAE34023;

XX 01-FEB-2002 (first entry)

DT Maize ASR protein.

DE Maize; ASR; abscisic acid-water stress-ripening-induced protein;

XX water stress resistance; crop plant.

XX Zea mays.

OS WO200183756-A1.

PN 08-NOV-2001.

XX 24-APR-2001; 2001WO-FR01252.

XX 28-APR-2000; 2000FR-0005534.

XX (BIOG-) BIOGENMA.

XX Zivvy M, Perez P;

PI WPI; 2002-041496/05.

DR N-PSDB; ABA01409.

XX Plant with improved resistance to water stress, contains sense or

PT antisense sequence encoding abscisic acid-water stress-ripening-induced

PT protein -

XX Claim 3; Page 32; 42pp; French.

XX The present invention relates to a method for the preparation of a plant

CC containing an altered amount of ASR protein (abscisic acid-water

CC stress-ripening-induced protein). Plants with altered expression of ASR

CC have increased resistance to water stress, relative to a non-transformed

CC plant. The method is especially applied to crop plants such as maize

CC (especially, wheat, rape, sunflower and peas to increase resistance to

CC water stress. The present sequence is a protein sequence for maize ASR

CC which was used to illustrate the present invention.

XX Sequence 140 AA;

SQ Query Match 80.0%; Score 36; DB 23; Length 140;

Best Local Similarity 83.3%; Pred. No. 34;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 HLYHHK 6
Db      10 HLFPHK 15

RESULT 11
AAU87835
ID      AAU87835 standard; Protein; 210 AA.
XX
AC
XX
XX
XX      05-JUN-2002 (first entry)
XX
DE      T. aureum 7091 partial elongase.
XX
KW      Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW      transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW      pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW      acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW      inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW      inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW      vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
OS      Thraustochytrium aureum.
XX
XX      WO200208401-A2.
XX
XX      31-JAN-2002.
XX
XX      24-JUL-2001; 2001WO-US23259.
XX
XX      24-JUL-2000; 2000US-0624670.
XX
XX      11-JUL-2001; 2001US-0903456.
XX
XX      (ABBO ) ABBOTT LAB.
XX
XX      Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
XX      Thurmond J, Pereira SL;
XX      WPI; 2002-172011/22.
XX      N-PSDB; ABK46404.
XX
XX      Isolated nucleic acid sequences encoding elongase proteins, useful in
XX      the production of polyunsaturated fatty acids, e.g. arachidonic acid,
XX      which can then be used nutritional compositions and pharmaceutical
XX      compositions (for treating AIDS) -
XX      Example 21; Fig 66; 271pp; English.
XX
XX      The invention relates to isolated nucleic acid sequences encoding
XX      elongase proteins and the elongase proteins themselves. Also
XX      included are a purified polypeptide which elongates polyunsaturated
XX      fatty acids and has at least 30% amino acid similarity to the amino acid
XX      sequence to an elongase protein, a vector comprising the nucleic
XX      acid, a plant cell (or tissue or whole plant) comprising the vector and
XX      expressing the nucleic acid, a plant oil or acid expressed by the
XX      transgenic plant, and a transgenic non-human animal expressing elongase
XX      in its fluid. The nucleic acids and elongase proteins may be used
XX      directly or indirectly in the production of polyunsaturated fatty acids
XX      (pufa), for e.g. arachidonic acid, which can then be used nutritional
XX      compositions, pharmaceutical compositions, cosmetics and animal feeds.
XX      The pharmaceutical compositions may be used in the treatment of
XX      restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
XX      multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
XX      urinary tract stones, malignant cancer, cachexia associated with cancer,
XX      eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
XX      psoriasis. They are also useful for inhibiting platelet aggregation,
XX      inducing vasodilation, reducing cholesterol, reducing or preventing
XX      gastrointestinal bleeding and side effects of non-steroidal
XX      anti-inflammatory drugs. The present sequence represents an elongase
XX      of the invention.

```

```

SQ      Sequence 210 AA;
Query Match      80.0%; Score 36; DB 23; Length 210;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHHK 7
Db      167 HVIHHAT 173

RESULT 12
AAU87836
ID      AAU87836 standard; Protein; 272 AA.
XX
XX
AC      AAU87836;
XX
XX      05-JUN-2002 (first entry)
XX
DE      T. aureum 7091 elongase TEL01 from plasmid pRAT-4-A1.
XX
KW      Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW      transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW      pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW      acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW      inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW      inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW      vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
XX      Thraustochytrium aureum.
XX
XX      WO200208401-A2.
XX
XX      31-JAN-2002.
XX
XX      24-JUL-2001; 2001WO-US23259.
XX
XX      24-JUL-2000; 2000US-0624670.
XX
XX      11-JUL-2001; 2001US-0903456.
XX
XX      (ABBO ) ABBOTT LAB.
XX
XX      Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
XX      Thurmond J, Pereira SL;
XX      WPI; 2002-172011/22.
XX      N-PSDB; ABK46405.
XX
XX      Isolated nucleic acid sequences encoding elongase proteins, useful in
XX      the production of polyunsaturated fatty acids, e.g. arachidonic acid,
XX      which can then be used nutritional compositions and pharmaceutical
XX      compositions (for treating AIDS) -
XX      Example 21; Fig 74; 271pp; English.
XX
XX      The invention relates to isolated nucleic acid sequences encoding
XX      elongase proteins and the elongase proteins themselves. Also
XX      included are a purified polypeptide which elongates polyunsaturated
XX      fatty acids and has at least 30% amino acid similarity to the amino acid
XX      sequence to an elongase protein, a vector comprising the nucleic
XX      acid, a plant cell (or tissue or whole plant) comprising the vector and
XX      expressing the nucleic acid, a plant oil or acid expressed by the
XX      transgenic plant, and a transgenic non-human animal expressing elongase
XX      in its fluid. The nucleic acids and elongase proteins may be used
XX      directly or indirectly in the production of polyunsaturated fatty acids
XX      (pufa), for e.g. arachidonic acid, which can then be used nutritional
XX      compositions, pharmaceutical compositions, cosmetics and animal feeds.
XX      The pharmaceutical compositions may be used in the treatment of
XX      restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
XX      multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
XX      urinary tract stones, malignant cancer, cachexia associated with cancer,
XX      eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
XX      psoriasis. They are also useful for inhibiting platelet aggregation,
XX      inducing vasodilation, reducing cholesterol, reducing or preventing
XX      gastrointestinal bleeding and side effects of non-steroidal
XX      anti-inflammatory drugs. The present sequence represents an elongase
XX      of the invention.

```


CC inducing vasodilation, reducing cholesterol, reducing or preventing
 CC gastrointestinal bleeding and side effects of non-steroidal
 CC anti-inflammatory drugs. The present sequence represents an elongase
 CC of the invention.
 XX
 SQ Sequence 272 AA;
 Query Match 80.0%; Score 36; DB 23; Length 272;
 Best Local Similarity 71.4%; Pred. No. 70;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HLYHHKT 7
 Db 152 HLYHHAT 158
 RESULT 13
 AAU87837
 ID AAU87837 standard; Protein; 272 AA.
 AC AAU87837;
 DT 05-JUN-2002 (first entry)
 DE T. aureum 7091 elongase TELO1 from plasmid pRAT-4-A2.
 XX
 KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
 KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
 KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
 KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
 KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
 KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
 KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
 XX
 OS Thraustochytrium aureum.
 XX WO200208401-A2.
 PN 31-JAN-2002.
 PD 24-JUL-2001; 2001WO-US23259.
 PF 24-JUL-2000; 2000US-0624670.
 PR 11-JUL-2001; 2001US-0903456.
 XX (ABBO) ABBOTT LAB.
 PA Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
 PI Thurmond J, Pereira SL;
 DR WPI; 2002-172011/22.
 XX N-PSDB; ABK46406.
 PT Isolated nucleic acid sequences encoding elongase proteins, useful in
 PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
 PT which can then be used nutritional compositions and pharmaceutical
 PT compositions (for treating AIDS) -
 XX Example 21; Fig 75; 271pp; English.
 PS
 CC The invention relates to isolated nucleic acid sequences encoding
 CC elongase proteins and the elongase proteins themselves. Also
 CC included are a purified polypeptide which elongates polyunsaturated
 CC fatty acids and has at least 30% amino acid similarity to the amino acid
 CC sequence to an elongase protein, a vector comprising the nucleic
 CC acid, a plant cell (or tissue or whole plant) comprising the vector and
 CC expressing the nucleic acid, a plant oil or acid expressed by the
 CC transgenic plant, and a transgenic non-human animal expressing elongase
 CC in its fluid. The nucleic acids and elongase proteins may be used
 CC directly or indirectly in the production of polyunsaturated fatty acids
 CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
 CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
 CC The pharmaceutical compositions may be used in the treatment of

CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
 CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
 CC urinary tract stones, malignant cancer, cachexia associated with cancer,
 CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
 CC psoriasis. They are also useful for inhibiting platelet aggregation,
 CC inducing vasodilation, reducing cholesterol, reducing or preventing
 CC gastrointestinal bleeding and side effects of non-steroidal
 CC anti-inflammatory drugs. The present sequence represents an elongase
 CC of the invention.
 XX
 SQ Sequence 272 AA;
 Query Match 80.0%; Score 36; DB 23; Length 272;
 Best Local Similarity 71.4%; Pred. No. 70;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HLYHHKT 7
 Db 152 HLYHHAT 158
 RESULT 14
 AAU87839
 ID AAU87839 standard; Protein; 272 AA.
 AC AAU87839;
 DT 05-JUN-2002 (first entry)
 DE T. aureum 7091 elongase TELO1 from plasmid pRAT-4-A4.
 XX
 KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
 KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
 KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
 KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
 KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
 KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
 KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
 XX
 OS Thraustochytrium aureum.
 XX WO200208401-A2.
 PN 31-JAN-2002.
 PD 24-JUL-2001; 2001WO-US23259.
 PF 24-JUL-2000; 2000US-0624670.
 PR 11-JUL-2001; 2001US-0903456.
 XX (ABBO) ABBOTT LAB.
 PA Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
 PI Thurmond J, Pereira SL;
 DR WPI; 2002-172011/22.
 XX N-PSDB; ABK46408.
 PT Isolated nucleic acid sequences encoding elongase proteins, useful in
 PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
 PT which can then be used nutritional compositions and pharmaceutical
 PT compositions (for treating AIDS) -
 XX Example 21; Fig 77; 271pp; English.
 PS
 CC The invention relates to isolated nucleic acid sequences encoding
 CC elongase proteins and the elongase proteins themselves. Also
 CC included are a purified polypeptide which elongates polyunsaturated
 CC fatty acids and has at least 30% amino acid similarity to the amino acid
 CC sequence to an elongase protein, a vector comprising the nucleic
 CC acid, a plant cell (or tissue or whole plant) comprising the vector and
 CC expressing the nucleic acid, a plant oil or acid expressed by the
 CC transgenic plant, and a transgenic non-human animal expressing elongase

CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX
SQ Sequence 272 AA;
Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 152 HVYHHAT 158
|:|||||
AAU87841 standard; Protein; 272 AA.
AAU87841;
XX
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 elongase TELO1 from plasmid pRAF-4-A6.
XX
XX Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
XX Thraustochytrium aureum.
XX
XX WO200208401-A2.
XX
XX 31-JAN-2002.
XX
XX 24-JUL-2001; 2001WO-US23259.
XX
XX 24-JUL-2000; 2000US-0624670.
XX
XX 11-JUL-2001; 2001US-0903456.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
XX WPI: 2002-172011/22.
XX
XX N-PSDB; ABK46371.
XX
XX Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
XX Example 21; Fig 79; 271pp; English.
PS
XX The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated

CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase
CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX
SQ Sequence 272 AA;
Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 152 HVYHHAT 158
|:|||||
Search completed: February 11, 2004, 17:03:03
Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-10

Perfect score: 45

Sequence: 1 HLYHHK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/ECTUS.COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	80.0	110	1	US-08-359-696-2
2	36	80.0	283	4	US-09-145-828A-22
3	35	77.8	349	4	US-09-011-769A-47
4	35	77.8	349	4	US-09-011-769A-60
5	35	77.8	349	4	US-09-011-769A-64
6	34	75.6	272	4	US-09-145-828A-18
7	33	73.3	458	5	PCT-US96-00994-4
8	33	73.3	855	4	US-08-890-865A-10
9	33	73.3	913	4	US-08-971-089-4
10	32	71.1	157	4	US-09-252-991A-27738
11	32	71.1	237	4	US-09-252-991A-18630
12	32	71.1	272	4	US-09-177-419C-2
13	32	71.1	298	4	US-09-177-419C-4
14	32	71.1	402	4	US-09-252-991A-26501
15	32	71.1	481	4	US-09-252-991A-23317
16	32	71.1	517	4	US-09-134-001C-3188
17	32	71.1	616	4	US-09-252-991A-19164
18	32	71.1	969	2	US-08-284-941-2
19	32	71.1	969	2	US-08-441-642-2
20	32	71.1	969	3	US-09-236-503-2
21	32	71.1	969	5	PCT-US93-02147A-2
22	31	68.9	147	4	US-09-145-828A-20
23	31	68.9	159	4	US-09-205-258-270
24	31	68.9	327	4	US-09-247-155-94
25	31	68.9	327	4	US-09-599-360B-12
26	31	68.9	348	4	US-09-252-991A-31001
27	31	68.9	354	4	US-09-107-532A-4236

ALIGNMENTS

RESULT 1
US-08-359-696-2
; Sequence 2, Application US/08359696
; Patent No. 5633440
; GENERAL INFORMATION:
; APPLICANT: DUNSMUIR, Pamela
; APPLICANT: STOTT, Jamie S.
; TITLE OF INVENTION: P-119 PROMOTERS AND THEIR USES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/359,696
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 12176-004800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-359-696-2

Query Match 80.0%; Score 36; DB 1; Length 110;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HLYHHK 6
|||
Db 10 HLFHHK 15

RESULT 2

Sequence 5312, Ap
Sequence 30946, A
Sequence 3358, Ap
Sequence 30001, A
Sequence 12, Appl
Sequence 28939, A
Sequence 38, Appl
Sequence 293, Appl
Sequence 4929, Ap
Sequence 95, Appl
Sequence 2, Appl
Sequence 39, Appl
Sequence 6165, Ap
Sequence 20, Appl
Sequence 42, Appl
Sequence 40, Appl
Sequence 4, Appl

28 31 68.9 427 4 US-09-328-352-5312
29 31 68.9 465 4 US-09-252-991A-30946
30 31 68.9 527 4 US-09-134-001C-3358
31 31 68.9 589 4 US-09-252-991A-30001
32 31 68.9 731 4 US-09-708-426-12
33 31 68.9 787 4 US-09-252-991A-28939
34 30 66.7 72 3 US-09-042-071-38
35 30 66.7 88 4 US-08-311-731A-293
36 30 66.7 103 4 US-09-328-352-4929
37 30 66.7 115 4 US-09-996-243-95
38 30 66.7 124 3 US-08-881-450A-2
39 30 66.7 124 3 US-09-042-071-39
40 30 66.7 163 4 US-09-328-352-6165
41 30 66.7 174 4 US-09-544-716-20
42 30 66.7 178 3 US-09-042-071-42
43 30 66.7 195 3 US-09-042-071-40
44 30 66.7 207 3 US-09-013-861-4
45 30 66.7 207 4 US-09-612-473-4

```

US-09-145-828A-22
; Sequence 22, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.O1
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (282)... (282)
; OTHER INFORMATION: Xaa = Unknown or other at position 282
US-09-145-828A-22

Query Match      80.0%; Score 36; DB 4; Length 283;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHKT 7
Db      125 HLYHAT 131

RESULT 3
US-09-011-769A-47
; Sequence 47, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; APPLICANT: BLAKEY, David C.
; APPLICANT: DAVIES, David H.
; APPLICANT: HENNAM, John F.
; APPLICANT: HENNEQUIN, Laurent F.A.
; APPLICANT: MARSHAM, Peter R.
; APPLICANT: DOWELL, Robert I.
; APPLICANT: MARSHAM, Peter R.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996
; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-011-769A-60

Query Match      77.8%; Score 35; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HLYHH 5
Db      327 HLYHH 331

RESULT 4
US-09-011-769A-60
; Sequence 60, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; APPLICANT: BLAKEY, David C.
; APPLICANT: DAVIES, David H.
; APPLICANT: HENNAM, John F.
; APPLICANT: HENNEQUIN, Laurent F.A.
; APPLICANT: MARSHAM, Peter R.
; APPLICANT: DOWELL, Robert I.
; APPLICANT: MARSHAM, Peter R.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996
; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-011-769A-60

Query Match      77.8%; Score 35; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HLYHH 5
Db      327 HLYHH 331

```

```

US-09-145-828A-22
; Sequence 22, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.O1
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (282)... (282)
; OTHER INFORMATION: Xaa = Unknown or other at position 282
US-09-145-828A-22

Query Match      80.0%; Score 36; DB 4; Length 283;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHKT 7
Db      125 HLYHAT 131

RESULT 3
US-09-011-769A-47
; Sequence 47, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; APPLICANT: BLAKEY, David C.
; APPLICANT: DAVIES, David H.
; APPLICANT: HENNAM, John F.
; APPLICANT: HENNEQUIN, Laurent F.A.
; APPLICANT: MARSHAM, Peter R.
; APPLICANT: DOWELL, Robert I.
; APPLICANT: MARSHAM, Peter R.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996
; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-011-769A-60

Query Match      77.8%; Score 35; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HLYHH 5
Db      327 HLYHH 331

```


;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham LLP
;; STREET: 1185 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: US
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/890,865A
;; FILING DATE: 10-JUL-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: White, John P
;; REGISTRATION NUMBER: 28,678
;; REFERENCE/DOCKET NUMBER: 0575/54249
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)278-0400
;; TELEFAX: (212)391-0526
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 855 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Protein
;; US-08-890-865A-10

Query Match 73.3%; Score 33; DB 4; Length 855;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
:|||||
Db 549 NLYHHK 554

RESULT 9
US-08-971-089-4
; Sequence 4, Application US/08971089
; Patent No. 6376174
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; APPLICANT: Scoles, Daniel R.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING
; TITLE OF INVENTION: SCHWANNOMIN-BINDING-PROTEINS AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,089
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,987
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramos, Robert T.
; REGISTRATION NUMBER: 37,915

;; REFERENCE/DOCKET NUMBER: P-CE 2862
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619)535-9001
;; TELEFAX: (619)535-8949
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 913 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-971-089-4

Query Match 73.3%; Score 33; DB 4; Length 913;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
:|||||
Db 572 HLYHH 576

RESULT 10
US-09-252-991A-27738
; Sequence 27738, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27738
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27738

Query Match 71.1%; Score 32; DB 4; Length 157;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
:|||||
Db 37 HLYHHK 42

RESULT 11
US-09-252-991A-18630
; Sequence 18630, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18630
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18630

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 26501
 LENGTH: 402
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26501

Query Match 71.1%; Score 32; DB 4; Length 402;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 |||||
 Db 300 HLLHHR 305

RESULT 15

US-09-252-991A-23317
 Sequence 23317, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 23317
 LENGTH: 481
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-23317

Query Match 71.1%; Score 32; DB 4; Length 481;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 |||||
 Db 74 HLLHHR 79

Search completed: February 11, 2004, 17:13:39
 Job time: 11.4167 secs

Query Match 71.1%; Score 32; DB 4; Length 237;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 |||||
 Db 213 HLLHHR 218

RESULT 12

US-09-177-419C-2
 Sequence 2, Application US/09177419C
 Patent No. 6562609
 GENERAL INFORMATION:
 APPLICANT: Russel, David W
 TITLE OF INVENTION: Cholesterol 25-Hydroxylase
 FILE REFERENCE: UTSD1370
 CURRENT APPLICATION NUMBER: US/09/177,419C
 CURRENT FILING DATE: 1998-10-22
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 272
 TYPE: PRT
 ORGANISM: human
 US-09-177-419C-2

Query Match 71.1%; Score 32; DB 4; Length 272;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 |||||
 Db 143 HLLHHR 148

RESULT 13

US-09-177-419C-4
 Sequence 4, Application US/09177419C
 Patent No. 6562609
 GENERAL INFORMATION:
 APPLICANT: Russel, David W
 TITLE OF INVENTION: Cholesterol 25-Hydroxylase
 FILE REFERENCE: UTSD1370
 CURRENT APPLICATION NUMBER: US/09/177,419C
 CURRENT FILING DATE: 1998-10-22
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 298
 TYPE: PRT
 ORGANISM: Mouse
 US-09-177-419C-4

Query Match 71.1%; Score 32; DB 4; Length 298;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 |||||
 Db 143 HLLHHR 148

RESULT 14

US-09-252-991A-26501
 Sequence 26501, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-10

Perfect score: 45

Sequence: 1 HLYHHKT 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	7	10	US-09-901-187B-10
2	36	80.0	210	10	Sequence 10, Appl
3	36	80.0	210	12	Sequence 68, Appl
4	36	80.0	272	10	Sequence 84, Appl
5	36	80.0	272	10	Sequence 75, Appl
6	36	80.0	272	10	Sequence 76, Appl
7	36	80.0	272	10	Sequence 78, Appl
8	36	80.0	272	10	Sequence 80, Appl
9	36	80.0	272	10	Sequence 81, Appl
10	36	80.0	272	12	Sequence 75, Appl
11	36	80.0	272	12	Sequence 76, Appl
12	36	80.0	272	12	Sequence 78, Appl
13	36	80.0	272	12	Sequence 80, Appl
14	36	80.0	272	12	Sequence 81, Appl
15	36	80.0	283	10	Sequence 29, Appl
16	36	80.0	283	10	Sequence 88, Appl

16	36	80.0	283	12	US-10-156-911-29
17	36	80.0	283	12	US-10-156-911-88
18	36	80.0	283	12	US-10-408-736-26
19	36	80.0	288	10	US-09-903-456-84
20	36	80.0	288	12	US-10-156-911-84
21	36	80.0	292	10	US-09-903-456-82
22	36	80.0	292	12	US-10-156-911-82
23	36	80.0	295	10	US-09-903-456-86
24	36	80.0	295	12	US-10-156-911-86
25	36	80.0	299	10	US-09-903-456-86
26	36	80.0	299	12	US-10-156-911-66
27	36	80.0	327	12	US-10-108-260A-3982
28	36	80.0	750	12	US-10-120-801-10
29	36	80.0	851	12	US-10-108-260A-4227
30	35	77.8	76	12	US-09-864-408A-6578
31	35	77.8	119	12	US-09-864-408A-7798
32	35	77.8	140	12	US-10-094-743-3159
33	35	77.8	291	12	US-10-374-780A-302
34	35	77.8	377	12	US-10-287-274-429
35	35	77.8	420	12	US-10-369-493-3766
36	35	77.8	613	12	US-10-369-493-2509
37	35	77.8	667	12	US-10-369-493-23387
38	35	77.8	668	9	US-09-720-989-20
39	35	77.8	1591	9	US-09-864-761-37952
40	34	75.6	62	9	US-09-864-761-35731
41	34	75.6	209	12	US-10-104-047-3213
42	34	75.6	216	12	US-09-833-245-1114
43	34	75.6	216	12	US-09-833-245-1115
44	34	75.6	236	12	US-09-976-782-106
45	34	75.6	263	15	US-10-050-704-171

ALIGNMENTS

RESULT 1
US-09-901-187B-10
Sequence 10, Application US/09901187B
Patent No. US20020151464A1
GENERAL INFORMATION:
APPLICANT: Panacea Pharmaceuticals, Inc.
APPLICANT: Wolozin, Benjamin
APPLICANT: Osetrova-Golts, Natalie
APPLICANT: Lebowitz, Michael S.
TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
TITLE OF INVENTION: Alpha-Synuclein Diseases
FILE REFERENCE: PAN01/002US
CURRENT APPLICATION NUMBER: US/09/901.187B
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/217,319
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/279,199
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens
US-09-901-187B-10

Query Match 100.0%; Score 45; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHKT 7

Db 1 HLYHHKT 7

RESULT 2
US-09-903-456-68
Sequence 68, Application US/09903456


```
Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-68

Query Match      80.0%; Score 36; DB 10; Length 210;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
Db      167 HVYHAT 173

RESULT 3
US-10-156-911-68
; Sequence 68, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-68

Query Match      80.0%; Score 36; DB 12; Length 210;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
Db      167 HVYHAT 173

Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-75

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
Db      152 HVYHAT 158

RESULT 4
US-09-903-456-75
; Sequence 75, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-75

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
Db      152 HVYHAT 158

RESULT 5
US-09-903-456-76
; Sequence 76, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-76

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
Db      152 HVYHAT 158
```

RESULT 6
 US-09-903-456-78
 ; Sequence 78, Application US/09903456
 ; Patent No. US20020138874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott Laboratories
 ; APPLICANT: Mukerji, Pradip
 ; APPLICANT: Leonard, Amanda Eun-Yeong
 ; APPLICANT: Huang, Yung-Sheng
 ; APPLICANT: Pereira, Suzette L.
 ; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
 ; FILE REFERENCE: 6407.US.P3
 ; CURRENT APPLICATION NUMBER: US/09/903,456
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 09/624,670
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: US 09/379,095
 ; PRIOR FILING DATE: 1999-08-23
 ; PRIOR APPLICATION NUMBER: US 09/145,828
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 78
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Thraustochytrium aureum
 US-09-903-456-78

Query Match 80.0%; Score 36; DB 10; Length 272;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
 | : | | | |
 Db 152 HVYHAT 158

RESULT 7
 US-09-903-456-80
 ; Sequence 80, Application US/09903456
 ; Patent No. US20020138874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott Laboratories
 ; APPLICANT: Mukerji, Pradip
 ; APPLICANT: Leonard, Amanda Eun-Yeong
 ; APPLICANT: Huang, Yung-Sheng
 ; APPLICANT: Pereira, Suzette L.
 ; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
 ; FILE REFERENCE: 6407.US.P3
 ; CURRENT APPLICATION NUMBER: US/09/903,456
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 09/624,670
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: US 09/379,095
 ; PRIOR FILING DATE: 1999-08-23
 ; PRIOR APPLICATION NUMBER: US 09/145,828
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 80
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Thraustochytrium aureum
 US-09-903-456-80

Query Match 80.0%; Score 36; DB 10; Length 272;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
 | : | | | |
 Db 152 HVYHAT 158

RESULT 8
 US-09-903-456-81
 ; Sequence 81, Application US/09903456
 ; Patent No. US20020138874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott Laboratories
 ; APPLICANT: Mukerji, Pradip
 ; APPLICANT: Leonard, Amanda Eun-Yeong
 ; APPLICANT: Huang, Yung-Sheng
 ; APPLICANT: Pereira, Suzette L.
 ; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
 ; FILE REFERENCE: 6407.US.P3
 ; CURRENT APPLICATION NUMBER: US/09/903,456
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 09/624,670
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: US 09/379,095
 ; PRIOR FILING DATE: 1999-08-23
 ; PRIOR APPLICATION NUMBER: US 09/145,828
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 81
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Thraustochytrium aureum
 US-09-903-456-81

Query Match 80.0%; Score 36; DB 10; Length 272;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
 | : | | | |
 Db 152 HVYHAT 158

RESULT 9
 US-10-156-911-75
 ; Sequence 75, Application US/10156911
 ; Publication No. US20030163845A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott Laboratories
 ; APPLICANT: Mukerji, Pradip
 ; APPLICANT: Leonard, Amanda Eun-Yeong
 ; APPLICANT: Huang, Yung-Sheng
 ; APPLICANT: Pereira, Suzette L.
 ; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
 ; FILE REFERENCE: 6407.US.P4
 ; CURRENT APPLICATION NUMBER: US/10/156,911
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: US 09/903,456
 ; PRIOR FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 09/624,670
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: US 09/379,095
 ; PRIOR FILING DATE: 1999-08-23
 ; PRIOR APPLICATION NUMBER: US 09/145,828
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 122
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 75
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Thraustochytrium aureum
 US-10-156-911-75

Query Match 80.0%; Score 36; DB 12; Length 272;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 HLYHHKT 7
      152 HYYHAT 158

Db

RESULT 10
US-10-156-911-76
; Sequence 76, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-76

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      152 HYYHAT 158

Db

RESULT 11
US-10-156-911-78
; Sequence 78, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-78

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      152 HYYHAT 158

Db

RESULT 12
US-10-156-911-80
; Sequence 80, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-80

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      152 HYYHAT 158

Db

RESULT 13
US-10-156-911-81
; Sequence 81, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-81
```

```

; SEQ ID NO 81
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-81

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
       |:||| |
Db      152 HVTYHAT 158

RESULT 14
US-09-903-456-29
; Sequence 29, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Potential Mammalian Elongase
; NAME/KEY: VARIANT
; LOCATION: (282)...(282)
; OTHER INFORMATION: Xaa = Unknown or Other at position 282
US-09-903-456-29

Query Match      80.0%; Score 36; DB 10; Length 283;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
       |:||| |
Db      134 HVTYHAT 140

Search completed: February 11, 2004, 17:54:13
Job time : 25.6667 secs

```

```

; SEQ ID NO 81
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-81

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
       |:||| |
Db      152 HVTYHAT 158

RESULT 14
US-09-903-456-29
; Sequence 29, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Potential Mammalian Elongase
; NAME/KEY: VARIANT
; LOCATION: (282)...(282)
; OTHER INFORMATION: Xaa = Unknown or Other at position 282
US-09-903-456-29

Query Match      80.0%; Score 36; DB 10; Length 283;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
       |:||| |
Db      125 HVTYHAT 131

RESULT 15
US-09-903-456-88
; Sequence 88, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24

```

OM protein - protein search, using sw model
Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHHKT 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: Piri:.*
2: Piri2:.*
3: Piri3:.*
4: Piri4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	86.7	305	H82702	small conductance
2	37	82.2	212	B83866	hypothetical prote
3	36	80.0	91	H86701	hypothetical prote
4	36	80.0	114	S37150	abz2 protein - com
5	36	80.0	115	T06588	abscisic stress ri
6	36	80.0	138	T02663	abscisic acid- and
7	36	80.0	169	T02081	ABA- and ripening-
8	36	80.0	259	T01976	hypothetical prote
9	36	80.0	325	T01975	hypothetical prote
10	35	77.8	242	G89633	glutamine transpor
11	35	77.8	292	T00823	wuschel protein -
12	35	77.8	367	D95196	oxidoreductase, Gf
13	35	77.8	367	A38063	conserved hypothet
14	35	77.8	367	A30634	probable oxidoredu
15	35	77.8	376	A12176	homocitrate syntha
16	35	77.8	377	S56505	hypothetical 41.9K
17	35	77.8	379	P35133	oxidoreductase, Gf
18	35	77.8	384	T49084	hypothetical prote
19	35	77.8	387	A82649	hypothetical prote
20	35	77.8	667	S56017	formate dehydrogen
21	35	77.8	982	T43699	DNA mismatch repai
22	35	77.8	3345	T13423	hypothetical prote
23	34	75.6	217	A12186	hypothetical prote
24	34	75.6	365	T26226	hypothetical prote
25	34	75.6	383	F84866	protein P20H11.5 [
26	34	75.6	515	T160800	calcitonin recepto
27	34	75.6	515	T149154	calcitonin recepto
28	34	75.6	529	T00677	hypothetical prote
29	34	75.6	616	JQ1441	hypothetical 67K p

30 B. subtilis protei
31 B. subtilis protei
32 conserved hypothet
33 transcription regu
34 protein from bacte
35 probable MADS-box
36 conserved hypothet
37 hypothethical prote
38 hypothethical prote
39 hypothethical prote
40 hypothethical prote
41 hypothethical prote
42 hypothethical prote
43 hypothethical prote
44 hypothethical prote
45 hypothethical prote

ALIGNMENTS

RESULT 1

H82702
small conductance mechanosensitive ion channel XF1258 [imported] - Xylella fastidiosa (C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: H82702
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82702
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-305 <SIM>
A:Cross-references: GB:A8003960; GB:A8003849; NID:99106242; PIDN:AAF84067.1; GSPDB:GN00
A:Experimental source: strain 945C
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver,
as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, H.M.F.N.; Madeira, H.M.F.N.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1258

Query Match 86.7%; Score 39; DB 2; Length 305;
Best Local Similarity 85.7%; Pred. No. 8.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7

|||||

282 HLYHHDT 288

RESULT 2

B83866
hypothetical protein BH1730 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83866
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: B83866
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <STO>
A;Cross-references: GB:BA000004; NID:g10174345; PIDN:BA805449.1; GSPDB:GN000
A;Experimental source: strain C-125
C;Genetics:

A;Gene: BHI730
C;Superfamily: Bacillus subtilis hypothetical protein yjpP

Query Match 82.2%; Score 37; DB 2; Length 212;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|||
Db 153 HLYHQKT 159

RESULT 3

H86701
Hypothetical protein ygaE [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: H86701

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se
A;Reference number: A86825; MUID:21235186; PMID:11337471

A;Accession: H86701
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <STO>
A;Cross-references: GB:AE005176; PID:g12723513; PIDN:AAK04714.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ygaE

Query Match 80.0%; Score 36; DB 2; Length 91;
Best Local Similarity 85.7%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|||
Db 56 HLCRHKT 62

RESULT 4

S37150
aar2 protein - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C;Accession: S37150

R;Amitai, H.; Scolnik, P.A.; Bar-Zvi, D.
submitted to the EMBL Data Library, September 1993

A;Reference number: S37150
A;Accession: S37150
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <AMI>
A;Cross-references: EMBL:X74907; NID:g400468; PID:g400469
C;Genetics:
A;Introns: 53/3

Query Match 80.0%; Score 36; DB 2; Length 114;
Best Local Similarity 83.3%; Pred. No. 9.9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 11 HLFHHK 16

RESULT 5

T06588
abscisic stress ripening protein 1 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T06588
R;Iusem, N.D.; Bartholomew, D.M.; Hitz, W.D.; Scolnik, P.A.
Plant Physiol. 102, 1353-1354, 1993

A;Title: Tomato (Lycopersicon esculentum) transcript induced by water deficit and ripen
A;Reference number: Z15778; MUID:9410533; PMID:8278555
A;Accession: T06588
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-115 <IUS>
A;Cross-references: EMBL:L08255; NID:g170375; PIDN:AAA34137.1; PID:g170376
A;Experimental source: strain Ailea Craig; fruit pericarp

Query Match 80.0%; Score 36; DB 2; Length 115;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 10 HLFHHK 15

RESULT 6

T02663
abscisic acid- and stress-induced protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: T02663
R;Vaidyanathan, R.; Kuruvilla, S.; Thomas, G.
submitted to the EMBL Data Library, December 1997
A;Description: Abscisic acid and stress inducible cDNA.
A;Reference number: Z14690
A;Accession: T02663
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-138 <VAL>
A;Cross-references: EMBL:AF039573; NID:g273153; PID:g2773154
A;Experimental source: cultivar Pokkali, Vytilla 1
C;Genetics:
A;Gene: ASR1

Query Match 80.0%; Score 36; DB 2; Length 139;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 9 HLFHHK 14

RESULT 7

T02081
ABA- and ripening-inducible-like protein - maize
C;Species: Zea mays (maize)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 29-Oct-1999
C;Accession: T02081
R;Arredondo-Peter, R.; Shearman, L.; Ji, L.; Klucas, R.V.
submitted to the EMBL Data Library, April 1994
A;Description: Nucleotide sequence of an ABA- and ripening-like cDNA isolated from corn
A;Reference number: Z14553
A;Accession: T02081
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-169 <ARR>
A;Cross-references: EMBL:U09276; NID:g551482; PIDN:AAA21866.1; PID:g551483
A;Experimental source: strain Golden Bantam; mesophyll

Query Match 80.0%; Score 36; DB 2; Length 169;

Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHK 6
DB 10 HLFHHK 15

RESULT 8

T01976
Hypothetical protein T9A4.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01976

R;Zidanic, M.; McQuerry, Y.; Smith, A.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of A. thaliana T9A4.
A:Reference number: Z14478

A:Accession: T01976
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-259 <ZID>
A:Cross-references: EMBL:AF096373; NID:G3695400; PID:G3695410
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 48/3; 79/3; 142/3; 194/3
A:Note: T9A4.10

Query Match 80.0%; Score 36; DB 2; Length 259;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
DB 74 HTFHHKT 80

RESULT 9

T01975
Hypothetical protein T9A4.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01975

R;Zidanic, M.; McQuerry, Y.; Smith, A.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of A. thaliana T9A4.
A:Reference number: Z14478

A:Accession: T01975
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <ZID>
A:Cross-references: EMBL:AF096373; NID:G3695400; PID:G3695413
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 62/3; 93/3; 156/3; 207/3
A:Note: T9A4.9

Query Match 80.0%; Score 36; DB 2; Length 325;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
DB 88 HTFHHKT 94

RESULT 10

G69633
Glutamine transport protein glnQ - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001

C:Accession: G69633
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
keuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dauchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: G69633
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-242 <KUN>
A:Cross-references: GB:Z99117; GB:Z99118; GB:AL009126; NID:G2635200; PID:NCBI4701.1; E
A:Experimental source: strain 168
C:Genetics:

A:Gene: glnQ
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; glutamine transport; nucleotide binding; P-loop
F17-212/Domain: ATP-binding cassette homology <ABC>
F134-41/Region: nucleotide-binding motif A (P-loop)
F157-161/Region: nucleotide-binding motif B

Query Match 77.8%; Score 35; DB 2; Length 242;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
DB 87 HLYPHKT 93

RESULT 11

T00829
wuschel protein - Arabidopsis thaliana
N:Alternate names: protein T13L16.3

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 30-Apr-1999
C:Accession: T00829
R;de la Bastide, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haber
McCombie, W.R.

submitted to the EMBL Data Library, January 1999
A:Description: A. thaliana BAC T13L16 from chromosome IV, top arm.
A:Reference number: Z14205
A:Accession: T00829
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-292
A:Cross-references: EMBL:AC003952; NID:G2708736; PID:G2708739
A:Experimental source: cultivar Columbia
C:Genetics:

A:Gene: wuschel
A:Map position: 2
A:Introns: 166/1; 195/3
A:Note: T13L16.3

Query Match 77.8%; Score 35; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
DB 152 HLYHH 156

RESULT 12

D95196
 oxidoreductase, Gfo/Idh/MocA family SP1686 [imported] - Streptococcus pneumoniae (strain D95196)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: D95196
 R:Tattelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf, E.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: D95196
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-367 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75765.1; PID:gl4973179; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP1686

Query Match 77.8%; Score 35; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
 |||||
 Db 170 HLYHH 174

RESULT 13

A98063
 conserved hypothetical protein spr1530 [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: A98063
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Eyer, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: A98063
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-367 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAL00334.1; PID:gl5459193; GSPDB:GN00174
 C:Genetics:
 A:Gene: spr1530

Query Match 77.8%; Score 35; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
 |||||
 Db 170 HLYHH 174

RESULT 14

AG0634
 probable oxidoreductase STY1170 [imported] - Salmonella enterica subsp. enterica serovar AG0634
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AG0634
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar AG0634
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG0634
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-367 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08258.1; PID:gl6502305; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1170

Query Match 77.8%; Score 35; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
 |||||
 Db 170 HLYHH 174

RESULT 15

AI2176
 homocitrate synthase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AI2176
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AI2176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAB74667.1; PID:gl7132062; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: nifV2
 C:Superfamily: hydroxymethylglutaryl-CoA lyase

Query Match 77.8%; Score 35; DB 2; Length 376;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
 |||||
 Db 241 HLYHH 245

Search completed: February 11, 2004, 17:11:52
 Job time : 15.3333 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-10

Perfect score: 45

Sequence: 1 HLYHHK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	80.0	40	1 UC12 MAIZE	P80618 zea mays (m
2	36	80.0	114	1 ASR2 LYCES	P37219 lycopersico
3	36	80.0	115	1 ASR1 LYCES	Q8655 lycopersico
4	35	77.8	367	1 YG86 STRPN	Q84728 streptococc
5	35	77.8	372	1 YJHC ECOLI	P39353 escherichia
6	35	77.8	376	1 NIV2 ANASF	P58637 anabaena sp
7	35	77.8	667	1 YVAF BACSU	P37519 bacillus su
8	35	77.8	935	1 AXN1 BRARE	P57094 brachydanio
9	35	77.8	954	1 K6P1 YARLI	P59680 varrowia li
10	35	77.8	982	1 M8H2 SCFPO	O74773 schizosacch
11	35	77.8	2483	1 PCX DROME	P8490 drosophila
12	34	75.6	270	1 ELO3 HUMAN	Q9NB03 homo sapien
13	34	75.6	271	1 ELO3 MOUSE	O35949 mus musculu
14	34	75.6	312	1 ELO4 MOUSE	Q9EGC4 mus musculu
15	34	75.6	314	1 ELO4 HUMAN	Q9GZK5 homo sapien
16	34	75.6	314	1 ELO4 MACPA	Q95K73 macaca fasc
17	34	75.6	326	1 RSF2 HUMAN	P50749 homo sapien
18	34	75.6	515	1 CALR MOUSE	Q60755 mus musculu
19	34	75.6	516	1 CALR RAT	P32214 rattus norv
20	33	73.3	259	1 MOB2 YEAST	P43563 saccharomyc
21	33	73.3	453	1 ZO6 XENLA	P18749 xenopus lae
22	33	73.3	802	1 ENAH MOUSE	Q03173 mus musculu
23	33	73.3	841	1 AXN CHICK	O42400 gallus gall
24	33	73.3	913	1 IF38 HUMAN	Q99613 homo sapien
25	33	73.3	2146	1 INSR DROME	P03208 drosophila
26	32	71.1	292	1 ELO2 MOUSE	Q9J1J4 mus musculu
27	32	71.1	296	1 ELO2 HUMAN	Q9NHB9 homo sapien
28	32	71.1	312	1 NRT2 CHICK	P55807 gallus gall
29	32	71.1	326	1 YAS5 METJA	Q50455 methanococc
30	32	71.1	331	1 Y011 NPVOP	Q65359 oryzaia pseu
31	32	71.1	360	1 GBF2 ARATH	P43775 arabidopsis
32	32	71.1	399	1 EFTU HELPU	Q94K19 helicobacte
33	32	71.1	435	1 YYS3 CABEL	P49191 caenorhabdi

Q8dlz2 synchococc
O02789 monodelphis
Q8y199 ralstonia s
Q9err8 mus musculu
Q9p2f9 homo sapien
Q9kbj4 bacillus ha
Q9htc5 p glucosam1
P28006 saccharomyc
Q9ygy0 xenopus lae
O63415 rattus norv
P29122 homo sapien
P20545 vaccinia vi

34 32 71.1 461 1 THIC SYNEL
35 32 71.1 518 1 TDT MONDO
36 32 71.1 580 1 SYQ_RALSO
37 32 71.1 581 1 Z319 MOUSE
38 32 71.1 582 1 Z319 HUMAN
39 32 71.1 595 1 THIC BACHD
40 32 71.1 610 1 GLMS PSERE
41 32 71.1 793 1 GAC1 YEAST
42 32 71.1 842 1 AXN XENLA
43 32 71.1 937 1 PAC1 RAT
44 32 71.1 969 1 PAC4 HUMAN
45 31 68.9 91 1 YVBE_VACCV

ALIGNMENTS

RESULT 1

UC12 MAIZE
ID UC12 MAIZE STANDARD; PRT; 40 AA.
AC P80618;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 237)
DE (Fragments).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID:4577;
RN [1]
RP SEQUENCE.
RC Tissue=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0, ITS MW IS: 21.7 kDa.
CC -!- SIMILARITY: TO THE TOMATO ABSICISIC STRESS RIPENING PROTEINS.
CC -!- CAUTION: THE ORDER OF THE PEPTIDES SHOWN IS UNCERTAIN.
DR Maize-2DPAGE; P80618; COLEOPTILE.
DR MaizeDB; 123942; -.
FT NON_TER 1 1
FT NON_CONS 15 16
FT NON_CONS 25 26
FT NON_TER 40 40
SQ SEQUENCE 40 AA; 4426 MW; 2E4B14B6B37B34C2 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 40;
Best Local Similarity 83.3%; Pred. No. 1.6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HLYHHK 6
DB 20 HLFPHK 25

RESULT 2

ASR2 LYCES
ID ASR2 LYCES STANDARD; PRT; 114 AA.
AC P37219;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Absciscic stress ripening protein 2.
GN ASR2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.

```

OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RX MEDLINE=95148753; PubMed=7846175;
RA Amitai-Zeigerson H., Scolnik P.A., Bar-Zvi D.;
RT "Genomic nucleotide sequence of tomato Asr2, a second member of the
RL stress/ripening-induced Asr1 gene family.";
RL Plant Physiol. 106:1699-1700(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74907; CAA52873.1; --
DR PIR; S37150;
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
FT DOMAIN 108 113 POLY-HIS.
FT DOMAIN 108 113 POLY-HIS.
SQ SEQUENCE 114 AA; 13020 MW; AE12FBBCD3631248 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 114;
Best Local Similarity 83.3%; Pred. No. 4.7; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1;

QY 1 HLYHHK 6
DB 11 HLFHHK 16

RESULT 3
ASR1 LYCES
ID ASR1 LYCES STANDARD; PRT; 115 AA.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Absciscic stress ripening protein 1.
GN ASR1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RX MEDLINE=94105353; PubMed=8278555;
RA Iusem N.D., Bartholomew D.M., Hitz W.D., Scolnik P.A.;
RT "Tomato (Lycopersicon esculentum) transcript induced by water deficit
RT and ripening.";
RL Plant Physiol. 102:1353-1354(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RA Gilad A., Amitai-Zeigerson H., Scolnik P.A., Bar-Zvi D.;
RT "Genomic nucleotide sequence of the tomato stress/ripening induced
RT Asr1 gene.";
RL (in) Plant Gene Register PGR97-042.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: BY LEAF WATER DEFICIT AND FRUIT RIPENING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08255; AAA34137.1; --
DR EMBL; U86130; AAB64185.1; --
DR PIR; T06588; T06588.
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
KW Nuclear protein.
FT DOMAIN 6 10 POLY-HIS.
FT DOMAIN 74 78 POLY-ALA.
SQ SEQUENCE 115 AA; 13130 MW; 0575CC69A73FA176 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 115;
Best Local Similarity 83.3%; Pred. No. 4.8; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1;

QY 1 HLYHHK 6
DB 10 HLFHHK 15

RESULT 4
YGB6 STRPN
ID YGB6 STRPN STANDARD; PRT; 367 AA.
AC Q54728;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase SP1686 (EC 1.1.1.1).
GN SP1686.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiolli S., Dickinson T., Hickey E.K.,
RA Holt I.B., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE OF 1-347 FROM N.A.
RC STRAIN=Serotype 6;
RX MEDLINE=96326329; PubMed=8759848;
RA Berry A.M., Lock R.A., Paton J.C.;
RT "Cloning and characterization of nanB, a second Streptococcus
RT pneumoniae neuraminidase gene, and purification of the NanB enzyme
RT from recombinant Escherichia coli.";
RL J. Bacteriol. 178:4854-4860(1996).
CC -!- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY. STRONG, TO E.COLI
CC YHFC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007461; AAC75765.1; --
DR EMBL; U43526; AAC44397.1; --
DR PIR; D95196; D95196.
DR TIGR; SP1686; --

```

DT InterPro; IPR000683; GFO IDH Moca_C.
DT InterPro; IPR004104; GFO IDH Moca_C.
DR Pfam; PF01408; GFO IDH Moca_C; 1.
DR Pfam; PF02894; GFO IDH Moca_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
FT CONFLICT 9 T -> A (IN REF. 2).
FT CONFLICT 156 P -> L (IN REF. 2).
SQ SEQUENCE 367 AA; 41095 MW; DE1D666A752325D CRC64;

Query Match 77.8%; Score 35; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 170 HLYHH 174
|||||
RESULT 5
YJHC ECOLI STANDARD; PRT; 372 AA.
AC P39353;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase yJHC (EC 1.1.1.1).
GN YJHC OR B4280.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes."
RL Nucleic Acids Res. 23:2105-2119 (1995).
CC -!- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14003; AAA97176.1; ALT INIT.
DR EMBL; AE000498; AAC77236.1; ALT INIT.
DR EcoGene; EG12545; YJhc.
DR InterPro; IPR000683; GFO IDH Moca_C.
DR InterPro; IPR004104; GFO IDH Moca_C.
DR Pfam; PF01408; GFO IDH Moca_C; 1.
DR Pfam; PF02894; GFO IDH Moca_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 372 AA; 41384 MW; 34B2E535EADC2ABE CRC64;

Query Match 77.8%; Score 35; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 170 HLYHH 174
|||||
RESULT 6
NIV2 ANASP STANDARD; PRT; 376 AA.
AC P58637;

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homocitrate synthase 2 (EC 2.3.3.14).
GN NIV2 OR ALR2968.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213 (2001).
CC -!- FUNCTION: THIS PROTEIN IS A FE-MO-COFACTOR BIOSYNTHETIC
CC COMPONENT.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + 2-oxoglutarate = 2-
CC hydroxybutane-1,2,4-tricarboxylate + CoA.
CC -!- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
CC synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003591; BAB74667.1; -.
DR PIR; A12176; A12176.
DR InterPro; IPR002034; AIPM/Hcit_synth.
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF06882; HMGL-like; 1.
DR PROSITE; PS00815; AIPM HOMOCIT SYNTH 1; 1.
DR PROSITE; PS00816; AIPM HOMOCIT SYNTH 2; 1.
KW Nitrogen fixation; Transferrase; Complete proteome.
SQ SEQUENCE 376 AA; 40936 MW; 343A804D990E4300 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 241 HLYHH 245
|||||
RESULT 7
YVAE BACSU STANDARD; PRT; 667 AA.
AC P37519;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Yvae.
GN YVAE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin."
RL DNA Res. 1:1-14 (1994).

DR InterPro; IPR000683; GFO IDH Moca_C.
DR InterPro; IPR004104; GFO IDH Moca_C.
DR Pfam; PF01408; GFO IDH Moca_C; 1.
DR Pfam; PF02894; GFO IDH Moca_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
FT CONFLICT 9 T -> A (IN REF. 2).
FT CONFLICT 156 P -> L (IN REF. 2).
SQ SEQUENCE 367 AA; 41095 MW; DE1D666A752325D CRC64;

Query Match 77.8%; Score 35; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 170 HLYHH 174
|||||
RESULT 5
YJHC ECOLI STANDARD; PRT; 372 AA.
AC P39353;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase yJHC (EC 1.1.1.1).
GN YJHC OR B4280.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes."
RL Nucleic Acids Res. 23:2105-2119 (1995).
CC -!- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14003; AAA97176.1; ALT INIT.
DR EMBL; AE000498; AAC77236.1; ALT INIT.
DR EcoGene; EG12545; YJhc.
DR InterPro; IPR000683; GFO IDH Moca_C.
DR InterPro; IPR004104; GFO IDH Moca_C.
DR Pfam; PF01408; GFO IDH Moca_C; 1.
DR Pfam; PF02894; GFO IDH Moca_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 372 AA; 41384 MW; 34B2E535EADC2ABE CRC64;

Query Match 77.8%; Score 35; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 170 HLYHH 174
|||||
RESULT 6
NIV2 ANASP STANDARD; PRT; 376 AA.
AC P58637;

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homocitrate synthase 2 (EC 2.3.3.14).
GN NIV2 OR ALR2968.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213 (2001).
CC -!- FUNCTION: THIS PROTEIN IS A FE-MO-COFACTOR BIOSYNTHETIC
CC COMPONENT.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + 2-oxoglutarate = 2-
CC hydroxybutane-1,2,4-tricarboxylate + CoA.
CC -!- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
CC synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003591; BAB74667.1; -.
DR PIR; A12176; A12176.
DR InterPro; IPR002034; AIPM/Hcit_synth.
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF06882; HMGL-like; 1.
DR PROSITE; PS00815; AIPM HOMOCIT SYNTH 1; 1.
DR PROSITE; PS00816; AIPM HOMOCIT SYNTH 2; 1.
KW Nitrogen fixation; Transferrase; Complete proteome.
SQ SEQUENCE 376 AA; 40936 MW; 343A804D990E4300 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 HLYHH 5
Db 241 HLYHH 245
|||||
RESULT 7
YVAE BACSU STANDARD; PRT; 667 AA.
AC P37519;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Yvae.
GN YVAE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin."
RL DNA Res. 1:1-14 (1994).

RESULT 8
AXNI BRANE STANDARD; PRT; 835 AA.
AC P57094;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
GN Axin 1 (Axis inhibition protein 1).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
CX NCBI_TaxID=7955;
RN [1].
RP MEDLINE FROM N.A.
RX MEDLINE=20171051; PubMed=10704853;
RA Shimizu T., Yamanaka Y., Ryu S.-L., Hashimoto H., Yabe T., Hirata T.,
RA Bae Y.-K., Hibi M., Hirano T.;
RT "Cooperative roles of Bozozok/Dharma and Nodal-related proteins in the
RL formation of the dorsal organizer in zebrafish.";
CC Mech. Dev. 91:293-303 (2000).
CC -!- FUNCTION: INHIBITOR OF THE WNT SIGNALING PATHWAY. DOWN-REGULATES
CC BETA-CATENIN. PROBABLY FACILITATES THE PHOSPHORYLATION OF BETA-
CC CATENIN AND APC BY GSK-3B (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B AND DEPHOSPHORYLATED BY
CC PP2A (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 RGS domain.
CC -!- SIMILARITY: Contains 1 DIX domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB032262; BAA92439.1; -.
DR HSP; P49799; 1AGR.
DR ZFIN; ZDB-GENE-000403-1; axin1.
DR InterPro; IPR001158; DIX.
DR InterPro; IPR000342; Regl_Gprotein.
DR Pfam; PF00778; DIX; 1.
DR Pfam; PF00615; RGS; 2.
DR PRINTS; PR01301; RGS-PROTEIN.
DR ProDom; PD003639; DIX; 1.
DR ProDom; PD001580; Regl_Gprotein; 1.
DR SMART; SM00315; RGS; 1.
DR PROSITE; PS00841; DIX; 1.
DR PROSITE; PS01332; RGS; 1.
KW Developmental protein; Wnt signaling pathway; Phosphorylation.
FT DOMAIN 92 214
FT DOMAIN 351 436
FT DOMAIN 437 512
FT DOMAIN 753 835
SQ SEQUENCE 835 AA; 94351 MW; 1C62FCF1F5937C87 CRC64;
Query Match 77.8%; Score 35; DB 1; Length 835;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHK 6
DB 538 HLYHHK 543
RESULT 9
K6PL_YARLI

```
ID K6P1_YARLI STANDARD; PRT; 954 AA.
AC P59680;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 6-phosphofructokinase alpha subunit [EC 2.7.1.11] (Phosphofructokinase
DE 1) (phosphofructokinase) (6PF-1-K alpha subunit).
GN PFK1.
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
SEQUENCE FROM N.A.
RA Flores C.L., Martinez-Costa O.H., Sanchez V., Aragon J.J.,
RA Gancedo C.;
RA "The phosphofructokinase from the yeast Yarrowia lipolytica.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -!- ENZYME REGULATION: Allosterically inhibited by ATP and activated
CC by AMP and fructose 2,6-bisphosphate (By similarity).
CC -!- PATHWAY: Key control step of glycolysis.
CC -!- SUBUNIT: Heterooctamer of 4 alpha and 4 beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphofructokinase family. Two domains
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY142710; FAN34943.1; -
CC DR PROSITE; P500433; PHOSPHOFRUCTOKINASE; 2.
CC KW Kinase; Transferase; Glycolysis; Repeat; Allosteric enzyme.
CC SEQUENCE 954 AA; 104049 MW; 33CD3B8C292F2450 CRC64;
CC
CC Query Match 77.8%; Score 35; DB 1; Length 954;
CC Best Local Similarity 85.7%; Pred. No. 64;
CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 HLYHHKT 7
CC DB 896 HLYAHKT 902
CC
CC RESULT 10
CC MSH2_SCHPO STANDARD; PRT; 982 AA.
ID MSH2_SCHPO STANDARD; PRT; 982 AA.
AC Q74773; O42950;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein msh2.
GN MSH2 OR SPBC19G7.01C OR SPBC24C6.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=99077962; PubMed=9858548;
RA Rudolph C., Kunz C., Parisi S., Lehmann E., Hartsuiker E.,
RA Partmann B., Kramer W., Kohli J., Fleck O.;
RT "The msh2 gene of Schizosaccharomyces pombe is involved in mismatch
RT repair, mating-type switching, and meiotic chromosome organization.";
RL Mol. Cell. Biol. 19:241-250(1999).
```

RESULT 11		
PCX_DROME	STANDARD;	PRT; 2483 AA.
ID	PCX_DROME	AC P18490;
DT	01-NOV-1990 (Rel. 16, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-OCT-1996 (Rel. 34, Last annotation update)	
DE	Pecanex protein.	
GN	PCX.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephyroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RN	[1]_SEQUENCE OF 1-545 FROM N.A.	
RP	STRAIN=Oregon-R;	
RC	Labonne S.G.;	
RL	Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE OF 546-2483 FROM N.A.	
RC	STRAIN=Oregon-R;	
EX	MEDLINE=90033754; PubMed=2478400;	
RA	Labonne S.G., Sunitha T., Mahowald A.P.;	
RT	"Molecular Genetics of pecanex, a maternal-effect neurogenic locus of Drosophila melanogaster that potentially encodes a large transmembrane protein.";	
RT	Dev. Biol. 136:1-16(1989).	
RL	FUNCTION: INVOLVED IN NEUROGENESIS.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).	
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	EMBL; M74329; AAA28747.1; -	
DR	EMBL; M25662; AAA28749.1; -	
DR	PIR; A37361; A37361.	
DR	FlyBase; FBgn003048; pcx.	
DR	Pfam; PF05041; Pecanex_C1.	
KW	Developmental protein; Neurogenesis; Transmembrane; Glycoprotein; Repeat.	
FT	TRANSMEM 750 771 POTENTIAL.	
FT	TRANSMEM 859 883 POTENTIAL.	
FT	TRANSMEM 1285 1307 POTENTIAL.	
FT	POLY-HIS. 39 51	
FT	DOMAIN 1231 1240 5 X 2 AA TANDem REPEATS OF G-T.	
FT	CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 117 117 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 711 711 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 902 902 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1118 1118 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1337 1337 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1350 1350 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1528 1528 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1688 1688 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 1786 1786 N-LINKED (GLCNAC. .) (POTENTIAL).	
FT	CARBOHYD 2315 2315 N-LINKED (GLCNAC. .) (POTENTIAL).	
SQ	SEQUENCE 2483 AA; 266639 MW; D376A3AB57112ABE CRC64;	
Query Match 77.8%; Score 35; DB 1; Length 2483;		
Best Local Similarity 83.3%; Pred. No. 1.7e+02;		
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		

RESULT 12		
ELO3_HUMAN	STANDARD;	PRT; 270 AA.
ID	ELO3_HUMAN	AC Q9HB03; Q8N180;
DT	28-FEB-2003 (Rel. 41, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Elongation of very long chain fatty acids protein 3 (Cold inducible glycoprotein of 30 kDa).	
GN	ELOVL3 OR CIG30.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]_SEQUENCE FROM N.A.	
RP	TISSUE=Skin;	
RC	MEDLINE=22388257; PubMed=12477932;	
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,	
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RL	[2]	
RP	SEQUENCE OF 35-270 FROM N.A.	
RC	Semina E.V., Murray J.C.;	
RA	"The CIG30/PTX3/GBF1 gene configuration is conserved between human and mouse genomes.";	
RT	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	
CC	-1- FUNCTION: May be involved in a membrane event related to cellular proliferation in brown adipose tissue. Could be implicated in synthesis of very long chain fatty acids and sphingolipids. May catalyze one or both of the reduction reaction in fatty acid elongation, i.e., conversion of beta-ketoacyl CoA to beta-hydroxyacyl CoA or reduction of trans-2-enoyl CoA to the saturated acyl CoA derivative (By similarity).	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (Potential).	
CC	-1- SIMILARITY: BELONGS TO THE ELO FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	EMBL; BC034344; AAH34344.1; -	
DR	EMBL; AF292387; AAG17875.1; -	
DR	Genew; HGNC.18047; ELOVL3.	
DR	InterPro; IPR002076; GNS1_SUR4.	
DR	Pfam; PF01151; ELO; 1.	

```

DR PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 29 POTENTIAL.
FT TRANSMEM 49
FT TRANSMEM 66
FT TRANSMEM 164 186
FT TRANSMEM 164 186
FT TRANSMEM 198 218
FT TRANSMEM 235 255
FT TRANSMEM 266 268
FT SITE
FT SEQUENCE 270 AA; 31500 MW; 0C6C8F1E7B5DE9B1 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 270;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HLYHKT 7
DB 145 HWYHST 151

RESULT 13
ELO3 MOUSE STANDARD; PRT; 271 AA.
AC 035943;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Elongation of very long chain fatty acids protein 3 (Cold inducible
DE Glycoprotein of 30 kDa) (CIN-2).
DE ELOV3 OR CIG30.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brown adipose tissue;
EX MEDLINE=98058971; PubMed=9395518;
RA Tvrdik P., Asadi A., Kozak L.P., Nedergaard J., Cannon B.,
RA Jacobson A.;
RT "Cig30, a mouse member of a novel membrane protein gene family, is
RT involved in the recruitment of brown adipose tissue."
RL J. Biol. Chem. 272:31738-31745(1997).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=99403085; PubMed=10473596;
RA Tvrdik P., Asadi A., Kozak L.P., Nuglozeh E., Parente F.,
RA Nedergaard J., Jacobson A.;
RT "Cig30 and Pitx3 genes are arranged in a partially overlapping
RT tail-to-tail array resulting in complementary transcripts."
RL J. Biol. Chem. 274:26387-26392(1999).
[3]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Colling F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.N., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be involved in a membrane event related to cellular
CC proliferation in brown adipose tissue. Could be implicated in
CC synthesis of very long chain fatty acids and sphingolipids. May
CC catalyze one or both of the reduction reaction in fatty acid
CC elongation, i.e., conversion of beta-ketoacyl CoA to beta-
CC hydroxyacyl CoA or reduction of trans-2-enoyl CoA to the saturated
CC acyl CoA derivative.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Readily detected in brown adipose tissue and
CC liver. Weak expression in kidney, white adipose tissue, heart and
CC skin. Not detected in lung, testis, muscle, spleen, brain, thymus
CC and intestine.
CC -!- INDUCTION: Elevated in brown adipose tissue in conditions of brown
CC fat recruitment, namely cold stress, perinatal development and
CC after diet-induced thermogenesis. A synergistic action of both
CC catecholamines and glucocorticoids is required for the induction.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J97107; AAC06127.1; -.
CC EMBL; AF054504; AAD51088.1; -.
CC EMBL; BC016483; AAH16488.1; -.
CC MGD; MGI:1195976; Elov13.
CC InterPro; IPR002076; GNS1_SUR4.
CC Pfam; PFO1151; ELO; 1.
CC PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum;
KW Glycoprotein.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 165 187 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 269 269 POTENTIAL.
FT SITE
FT CARBOHYD 6 6 N-LINKED (GLCNAC...) (POTENTIAL).
FT SEQUENCE 271 AA; 32060 MW; F7CA96199BE89401 CRC64;
SQ
Query Match 75.6%; Score 34; DB 1; Length 271;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HLYHKT 7
DB 146 HWYHST 152

RESULT 14
ELO4 MOUSE STANDARD; PRT; 312 AA.
AC 09EQC4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation of very long chain fatty acids protein 4.
DE ELOV4.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP

```

RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=20578755; PubMed=11138005;
RA Zhang K., Kniazeva M., Han M., Li W., Yu Z., Yang Z., Li Y.,
RA Metzger M.L., Allikmets R., Zack D.J., Kakuk L.E., Lagali P.S.,
RA Wong P.W., McDonald I.M., Sieving P.A., Figueroa D.J., Austin C.P.,
RA Gould R.J., Ayyagari R., Petrukhin K.;
RT "A 5-bp deletion in ELOVL4 is associated with two related forms of
RT autosomal dominant macular dystrophy.";
RL Nat. Genet. 27:89-93(2001).
CC -!- FUNCTION: Involved in the biosynthesis of very long chain fatty
CC acids. Seems to represent a photoreceptor-specific component of
CC the fatty acid elongation system residing on the endoplasmic
CC reticulum.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in the adult retina, exclusively in
CC photoreceptor cells.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF277093; AAG47667.1; -
CC MGD; MGI:1933331; Elov14.
CC InterPro: IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS0188; ELO; 1.
CC Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
CC TRANSMEM 42 62 POTENTIAL.
CC FT TRANSMEM 78 98 POTENTIAL.
CC FT TRANSMEM 165 185 POTENTIAL.
CC FT TRANSMEM 188 208 POTENTIAL.
CC FT TRANSMEM 246 266 POTENTIAL.
CC FT SITE 308 310 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
CC (POTENTIAL).
CC FT CARBOHYD 20 20 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT SEQUENCE 312 AA; 36520 MW; B04CD48024772132 CRC64;
SQ
Query Match 75.6%; Score 34; DB 1; Length 312;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHT 7
|:|:|:|
Db 158 HLYHHT 164
RESULT 15
ELO4_HUMAN STANDARD; PRT; 314 AA.
AC Q9GZK5; Q9H139;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation of very long chain fatty acids protein 4.
GN ELOVL4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT VAL-299.
RC TISSUE=Retina;
RX MEDLINE=20578755; PubMed=11138005;
RA Zhang K., Kniazeva M., Han M., Li W., Yu Z., Yang Z., Li Y.,
RA Metzger M.L., Allikmets R., Zack D.J., Kakuk L.E., Lagali P.S.,
RA Wong P.W., McDonald I.M., Sieving P.A., Figueroa D.J., Austin C.P.,
RA Gould R.J., Ayyagari R., Petrukhin K.;
RT "A 5-bp deletion in ELOVL4 is associated with two related forms of
RT autosomal dominant macular dystrophy.";
RL Nat. Genet. 27:89-93(2001).
CC -!- FUNCTION: Involved in the biosynthesis of very long chain fatty
CC acids. Seems to represent a photoreceptor-specific component of
CC the fatty acid elongation system residing on the endoplasmic
CC reticulum.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in the adult retina, exclusively in
CC photoreceptor cells.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF277093; AAG47667.1; -
CC MGD; MGI:1933331; Elov14.
CC InterPro: IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS0188; ELO; 1.
CC Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
CC TRANSMEM 42 62 POTENTIAL.
CC FT TRANSMEM 78 98 POTENTIAL.
CC FT TRANSMEM 165 185 POTENTIAL.
CC FT TRANSMEM 188 208 POTENTIAL.
CC FT TRANSMEM 246 266 POTENTIAL.
CC FT SITE 308 310 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
CC (POTENTIAL).
CC FT CARBOHYD 20 20 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT SEQUENCE 312 AA; 36520 MW; B04CD48024772132 CRC64;
SQ

RT "A 5-bp deletion in ELOVL4 is associated with two related forms of
RT autosomal dominant macular dystrophy.";
RL Nat. Genet. 27:89-93(2001).
CC -!- FUNCTION: Involved in the biosynthesis of very long chain fatty
CC acids. Seems to represent a photoreceptor-specific component of
CC the fatty acid elongation system residing on the endoplasmic
CC reticulum. May be implicated in docosahexaenoic acid (DHA)
CC biosynthesis, which requires dietary consumption of the essential
CC alpha-linolenic acid and a subsequent series of three elongation
CC steps. May be involved in one of these three elongation steps.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in the retina and at much lower
CC level in the brain.
CC -!- DISEASE: Defects in ELOVL4 are the cause of Stargardt disease 3
CC macular dystrophy (STGD3) and autosomal dominant macular dystrophy
CC (AMD). Both diseases are autosomal dominant inherited forms of
CC macular degeneration characterized by decreased visual acuity,
CC macular atrophy, and extensive fundus flecks.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF279654; AAG47669.1; -
CC EMBL; AF279649; AAG47669.1; JOINED.
CC EMBL; AF279650; AAG47669.1; JOINED.
CC EMBL; AF279651; AAG47669.1; JOINED.
CC EMBL; AF279652; AAG47669.1; JOINED.
CC EMBL; AF279653; AAG47669.1; JOINED.
CC EMBL; AF277094; AAG47668.1; -
CC EMBL; AY037298; AAK68639.1; -
CC EMBL; AK055277; BAB70895.1; -
CC EMBL; AL132875; CAC19496.1; -
CC Genew; HGNC:14415; ELOVL4.
CC MIM; 605512; -
CC MIM; 600110; -
CC GO; GO:0008020; F.G-protein coupled photoreceptor activity; NAS.
CC GO; GO:0006633; P.fatty acid biosynthesis; NAS.
CC InterPro: IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS0188; ELO; 1.
CC Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum;
CC Polymorphism; Stargardt disease; Vision.
KW

FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT SITE 310 312 POTENTIAL.
FT CARBOHYD 20 20 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
FT VARIANT 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 314 AA; 36829 MW; B2EBCE54D868E96E CRC64;
/FTid=VAR_012492.
M -> V.
Query Match 75.6%; Score 34; DB 1; Length 314;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|:|
DB 158 HVYHCT 164

Search completed: February 11, 2004, 17:04:14
Job time : 6.16667 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHHKT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:**

- 1: sp arChae:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_plage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	86.7	305	16 Q9PDX1	Q9PDX1 xylella fas
2	38	84.4	191	10 Q9FM60	Q9FM60 arabidopsis
3	37	82.2	212	16 Q9KC44	Q9KC44 bacillus ha
4	37	82.2	221	5 Q9VTF0	Q9VTF0 drosophila
5	37	82.2	222	5 Q81QF0	Q81QF0 drosophila
6	36	80.0	91	16 Q8CHV2	Q8CHV2 lactococcus
7	36	80.0	103	10 Q40165	Q40165 lycopersico
8	36	80.0	108	10 Q9ZRB5	Q9ZRB5 solanum tub
9	36	80.0	109	10 Q9ZRB6	Q9ZRB6 solanum tub
10	36	80.0	110	10 Q92575	Q92575 lycopersico
11	36	80.0	138	10 Q49149	Q49149 oryza sativ
12	36	80.0	142	10 Q948L3	Q948L3 saccharum o
13	36	80.0	149	10 Q94G23	Q94G23 vitis vinif
14	36	80.0	169	10 Q41730	Q41730 zea mays (m
15	36	80.0	218	11 Q9BHM5	Q9BHM5 mus musculu
16	36	80.0	259	10 Q82620	Q82620 arabidopsis

17	36	80.0	278	10 Q9SQV0	Q9SQV0 arabidopsis
18	36	80.0	282	13 Q8AX86	Q8AX86 brachydanio
19	36	80.0	298	10 Q9SQU9	Q9SQU9 arabidopsis
20	36	80.0	299	11 Q920L7	Q920L7 rattus norv
21	36	80.0	299	11 Q8R5D3	Q8R5D3 mus musculu
22	36	80.0	299	11 Q8BUE3	Q8BUE3 mus musculu
23	36	80.0	299	11 Q8BLG6	Q8BLG6 mus musculu
24	36	80.0	299	11 Q8BHI7	Q8BHI7 mus musculu
25	36	80.0	325	10 Q82619	Q82619 arabidopsis
26	36	80.0	327	4 Q8N8G4	Q8N8G4 homo sapien
27	36	80.0	350	5 Q9NE06	Q9NE06 leishmania
28	36	80.0	540	4 Q81YF3	Q81YF3 homo sapien
29	36	80.0	851	4 Q8N861	Q8N861 homo sapien
30	36	80.0	4832	5 Q81107	Q81107 plasmodium
31	35	77.8	140	4 Q96M18	Q96M18 homo sapien
32	35	77.8	242	16 Q34677	Q34677 bacillus su
33	35	77.8	277	5 Q9VC20	Q9VC20 drosophila
34	35	77.8	291	10 Q9SB92	Q9SB92 arabidopsis
35	35	77.8	292	10 Q9SL51	Q9SL51 arabidopsis
36	35	77.8	293	13 Q8AWE8	Q8AWE8 oreochromis
37	35	77.8	367	2 Q9F924	Q9F924 pasteurella
38	35	77.8	367	16 Q9CK74	Q9CK74 pasteurella
39	35	77.8	367	16 Q8Z7M7	Q8Z7M7 salmoneilla
40	35	77.8	367	16 Q8DN05	Q8DN05 streptococc
41	35	77.8	368	16 Q8ZQ31	Q8ZQ31 salmoneilla
42	35	77.8	379	2 Q9L4U9	Q9L4U9 anabaena va
43	35	77.8	379	16 Q97QA0	Q97QA0 streptococc
44	35	77.8	384	10 Q9SUZ9	Q9SUZ9 arabidopsis
45	35	77.8	387	16 Q9PCS4	Q9PCS4 xylella fas

ALIGNMENTS

RESULT 1

Q9PDX1 PRELIMINARY; PRT; 305 AA.

AC Q9PDX1; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Small conductance mechanosensitive ion channel.

GN XF1258
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Arya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper S.L., Kitajima J.P.,
RA Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Teai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
RL Nature 406:151-159(2000).
DR EMBL; AB003960; AAF84087.1; -;
DR InterPro; IPR006685; MSion_channel.
DR Pfam; PF00924; MS_channel; 1.
KW Complete proteome.
SQ SEQUENCE 305 AA; 33648 MW; 10F62B51496A2D53 CRC64;
Query Match 86.7%; Score 39; DB 16; Length 305;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
DB 282 HLYHHKT 288
RESULT 2
Q9FM60 PRELIMINARY; PRT; 191 AA.
AC Q9FM60;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Genomic DNA, chromosome 5, pl clone.MDF20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Columbia;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned pl and TAC clones";
RL DNA Res. 5:41-54(1998).
DR EMBL; AB009050; BAB09246.1; -;
SQ SEQUENCE 191 AA; 21713 MW; DEDF9AD52940E056 CRC64;
Query Match 84.4%; Score 38; DB 10; Length 191;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
DB 50 HLYHHKT 56
RESULT 3
Q9KC44 PRELIMINARY; PRT; 212 AA.
AC Q9KC44;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH1730.
GN BH1730.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512592; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
RT halodurans and genomic sequence comparison with *Bacillus subtilis*,"
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001513; BAB05449.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 212 AA; 24141 MW; E0A775F501BDD9B2 CRC64;
Query Match 82.2%; Score 37; DB 16; Length 212;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
DB 153 HLYHHKT 159
RESULT 4
Q9VTF0 PRELIMINARY; PRT; 221 AA.
AC Q9VTF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG11801 protein.
GN CG11801.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*,"

Science 287:2185-2195(2000).

[2] SEQUENCE FROM N.A.

RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brannon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson K., Pfeiffer B.,
 RA Pacle J.L., Paragay V., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Shapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RA "Sequencing of Drosophila melanogaster genome.";
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RA "Annotation of Drosophila melanogaster genome.";
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE FROM N.A.

RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[5] SEQUENCE FROM N.A.

RA FlyBase;
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003546; AAF50101.2; -;
 DR FlyBase; FBgn0036128; CG11801.
 DR InterPro; IPR002076; GNS1_SUR4.
 DR Pfam; PF01151; ELO; 1;
 DR PROSITE; PS01188; ELO; 1;
 DR PROSITE; 221 AA; 26345 MW; F0DF5FAS299365 CRC64;
 DR SEQUENCE 221 AA; 26345 MW; F0DF5FAS299365 CRC64;

Query Match 82.2%; Score 37; DB 5; Length 221;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HLYHKT 7
 Db 99 HVTHT 105

RESULT 5

Q81QF0 PRELIMINARY; PRT; 222 AA.

AC Q81QF0;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE CG32072-PA.
 GN CG32072.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cleverley S., Dahlke C., Davenport L.B., Davies S.M.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodak A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Nuzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Ye J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster.";
 RA Science 287:2185-2195(2000).

[2] SEQUENCE FROM N.A.

RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brannon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson K., Pfeiffer B.,
 RA Pacle J.L., Paragay V., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Shapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RA "Sequencing of Drosophila melanogaster genome.";
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RA "Annotation of Drosophila melanogaster genome.";
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE FROM N.A.

RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[5] SEQUENCE FROM N.A.

RA FlyBase;
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003546; AAF50101.2; -;
 DR FlyBase; FBgn0036128; CG11801.
 DR InterPro; IPR002076; GNS1_SUR4.
 DR Pfam; PF01151; ELO; 1;
 DR PROSITE; PS01188; ELO; 1;
 DR PROSITE; 221 AA; 26345 MW; F0DF5FAS299365 CRC64;
 DR SEQUENCE 221 AA; 26345 MW; F0DF5FAS299365 CRC64;

Query Match 82.2%; Score 37; DB 5; Length 221;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HLYHKT 7
 Db 99 HVTHT 105

RESULT 5

Q81QF0 PRELIMINARY; PRT; 222 AA.

AC Q81QF0;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE CG32072-PA.
 GN CG32072.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

Query Match      82.2%; Score 37; DB 5; Length 222;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 HLYHKT 7
DB 100 HLYHST 106

RESULT 6
Q9CHV2 PRELIMINARY; PRT; 91 AA.
AC Q9CHV2;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Unknown protein.
GN YGAE OR LL0616.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weisenbach J., Ehrlich S.B., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006294; AAK04714.1; -.
KW Complete proteome.
SQ SEQUENCE 91 AA; 10517 MW; A299E9C1682CSBFD CRC64;

Query Match      80.0%; Score 36; DB 16; Length 91;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
DB 56 HLYHKT 62

RESULT 7
Q40165 PRELIMINARY; PRT; 103 AA.
AC Q40165;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE ABA- and ripening-induced protein (fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94218393; PubMed=8165244;
RX Rossi M., Iusem N.D.;
RA "Tomato (Lycopersicon esculentum) genomic clone homologous to a gene
RT encoding an abscisic acid-induced protein."
RL Plant Physiol. 104:1073-1074(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=95352822; PubMed=7626782;
RX Rossi M., Iusem N.D.;
RA "Sequence of Asr2, a member of a gene family from Lycopersicon
RT esculentum encoding chromosomal proteins: homology to an intron of the
RT polygalacturonase gene."
RL DNA Seq. 5:225-227(1995).
DR EMBL; L20756; AAA99440.1; -.

```

```

DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
FT NON_TER 103 AA; 11631 MW; 926ED85B07660C6F CRC64;
SQ SEQUENCE 103 AA; 11631 MW; 926ED85B07660C6F CRC64;

Query Match      80.0%; Score 36; DB 10; Length 103;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHKT 6
DB 11 HLYHKT 16

RESULT 8
Q9ZRB5 PRELIMINARY; PRT; 108 AA.
ID Q9ZRB5;
AC Q9ZRB5;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE C121B protein.
GN C121B.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Saturna;
RX Schneider A., Salamini F., Gebhardt C.;
RA "Expression patterns and promoter activity of the cold regulated gene
RT c121A of potato."
RL Plant Physiol. 0:0-0(1996).
DR EMBL; U76611; AAD00255.1; -.
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
SQ SEQUENCE 108 AA; 12159 MW; 03CE2ECF8179BA30 CRC64;

Query Match      80.0%; Score 36; DB 10; Length 108;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHKT 6
DB 9 HLYHKT 14

RESULT 9
Q9ZRB6 PRELIMINARY; PRT; 109 AA.
ID Q9ZRB6;
AC Q9ZRB6;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE C121A protein.
GN C121A.
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Saturna;
RX Schneider A., Salamini F., Gebhardt C.;
RA "Expression patterns and promoter activity of the cold regulated gene
RT c121A of potato."
RL Plant Physiol. 0:0-0(1996).
DR EMBL; U76610; AAD00254.1; -.
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.

```

SQ SEQUENCE 109 AA; 12429 MW; 0981EGF4509F0135 CRC64;
Query Match 80.0%; Score 36; DB 10; Length 109;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
DB 9 HLFHHK 14

RESULT 10
Q82575 PRELIMINARY; PRT; 110 AA.
AC O82575;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Fruit-ripening protein.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridaceae; Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVRDC CL5915-93D4-1-0-3;
RA Wang Y.-C.; Heien H.-L.;
RT "Nucleotide sequence of a cDNA encoding a tomato fruit-ripening protein."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF093141; AAC61780.1; -;
DR InterPro; IPR003496; ABA WDS.
DR Pfam; PF02496; ABA WDS; 1.
SQ SEQUENCE 110 AA; 12555 MW; 467416108F74F363 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 110;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
DB 10 HLFHHK 15

RESULT 11
O49149 PRELIMINARY; PRT; 138 AA.
AC O49149;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Abscissic acid-and stress-inducible protein.
GN ASR1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Pokkali;
RA Vaidyanathan R.; Kuruvilla S.; Thomas G.;
RT "Characterization and expression pattern of an abscissic acid and osmotic stress responsive gene from rice."
RL Plant Sci. 140:25-36(1999).
DR EMBL; AF039573; AAB96681.1; -;
DR Gramene; O49149; -;
DR InterPro; IPR003496; ABA WDS.
DR Pfam; PF02496; ABA WDS; 1.
SQ SEQUENCE 138 AA; 15465 MW; 3E80C8E99C641D0E CRC64;

Query Match 80.0%; Score 36; DB 10; Length 138;

Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
DB 9 HLFHHK 14

RESULT 12
Q948L3 PRELIMINARY; PRT; 142 AA.
AC Q948L3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Drought inducible 22 kDa protein.
GN SODIP22.
OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M442-51;
RA Sugiharto B.; Ermawati N.; Mori H.; Aoki K.; Sakakibara K.Y.;
RA Yamaya T.; Sugiyama T.; Sakakibara H.;
RT "Identification and characterization of a gene encoding drought-inducible protein that localizes in the bundle sheath cell of sugarcane."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071694; BAB60268.1; -;
DR InterPro; IPR003496; ABA WDS.
DR Pfam; PF02496; ABA WDS; 1.
SQ SEQUENCE 142 AA; 15932 MW; 35DC5BC6D6A75A1 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 142;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
DB 9 HLFHHK 14

RESULT 13
Q94G23 PRELIMINARY; PRT; 149 AA.
AC Q94G23;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative transcription factor.
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RA Atanassova R.A.; Cakir B.; Gaillard C.; Delrot S.;
RT "A transcription factor binding to the promoter of a grape berry RT hexose transporter."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281656; AAK69513.1; -;
DR InterPro; IPR003496; ABA WDS.
DR Pfam; PF02496; ABA WDS; 1.
SQ SEQUENCE 149 AA; 16703 MW; D46C6B82953F66E5 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 149;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Job time : 27.5833 secs

OY 1 HLYHHK 6

DB 9 HLFHHK 14

RESULT 14
Q41730 PRELIMINARY; PRT; 169 AA.
AC Q41730;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE ABA- and ripening-inducible-like protein.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Arredondo-Peter R., Shearman L., Ji L., Klucas R.V.;
RT "Nucleotide sequence of an ABA- and ripening-like cDNA isolated from
RT corn roots."
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U09276; AA21866.1; -;
SQ SEQUENCE 169 AA; 18503 MW; 024AF2AA3D07B71F CRC64;

Query Match 80.0%; Score 36; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HLYHHK 6

DB 10 HLFHHK 15

RESULT 15
Q8BHM5 PRELIMINARY; PRT; 218 AA.
AC Q8BHM5;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Hypothetical FIVE Zn-finger.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL: AK049341; BAC33694.1; -;
DR EMBL: AK086031; BAC39597.1; -;
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 24303 MW; 03306ADFF63E85BB CRC64;

Query Match 80.0%; Score 36; DB 11; Length 218;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 HLYHHK 7

DB 186 HLYHHK 192

Search completed: February 11, 2004, 17:09:44

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)

Title: US-09-901-187C-11

perfect score: 44

Sequence: 1 THIHPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 segs. 158726573

Total number of hits satisfying chosen parameters

Minimum DB seg length: 0

Maximum DB seq length: 20000000000

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Maximum Match 100

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SIDSI/cgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Query			DB	ID	Description
			Match	Length	\$			
1	44	100.0	7	23	AAE14556		Human alpha-synuc	
2	43	97.7	148	24	ABP57617		S. murayamaensis A	
3	39	88.6	178	22	AAU19257		Human G protein-co	
4	39	88.6	248	22	ABG18105		Novel human digno	
5	39	88.6	301	22	ABG19187		Novel human digno	
6	39	88.6	301	22	ABG21623		Novel human digno	
7	39	88.6	322	22	ABG21623		Novel human digno	
8	39	88.6	334	22	ABG21635		Novel human digno	
9	39	88.6	383	22	ABG59684		Novel human digno	

10	39	88.6	393	22	ABG19170
11	39	88.6	402	14	AA041310
12	39	88.6	437	22	ABG02252
13	39	88.6	451	22	ABG19205
14	39	88.6	455	22	ABG10309
15	39	88.6	455	22	ABG21601
16	39	88.6	482	22	ABG05178
17	39	88.6	495	22	ABG19192
18	39	88.6	505	22	ABG21655
19	39	88.6	517	22	AAU31031
20	39	88.6	524	22	ABG02251
21	39	88.6	524	22	ABG19196
22	39	88.6	524	22	ABG23240
23	39	88.6	536	22	ABG29174
24	39	88.6	546	22	ABG02250
25	39	88.6	546	22	ABG19203
26	39	88.6	548	22	ABG1453
27	39	88.6	568	22	ABG28701
28	39	88.6	601	22	ABG02256
29	39	88.6	632	22	ABG17784
30	39	88.6	645	22	ABG02255
31	39	88.6	646	23	AAU12293
32	39	88.6	665	22	ABG04325
33	39	88.6	673	22	ABG02254
34	39	88.6	728	22	ABG19197
35	39	88.6	786	22	ABG02257
36	39	88.6	839	22	ABG22934
37	39	88.6	863	22	ABG19208
38	39	88.6	1067	22	ABB12481
39	39	88.6	1130	22	ABG02249
40	39	88.6	1455	22	ABG02259
41	39	88.6	1481	22	ABG13597
42	39	86.4	1779	22	ABH95562
43	38	86.4	1377	22	ABH64706
44	38	86.4	1878	19	AAW8170
45	38	86.4	1878	22	AAW40239

ALIGNMENTS

RESULT 1

AAE14556
ID AAE14556 standard; peptide; 7 AA.

AAE14556:

17-MAY-2002 (first entry)

Human alpha-synuclein aggregation inhibitor #11.

KW Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
 KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
 KW multiple system atrophy; Hallervorden-Spatz disease; human.

OS Homo sapiens.

XX PN WO20020482-A1.

4-XX

17-JAN-2002.



PF ~~06-JUL-2001~~; 2001WO-US21379.

PR 07-JUL-2000; 2000US-217319P.

PR 28-MAR-2001; 2001US-279199P.

✓ DATE / DATE TIME TNC

PA (PANA-) PANACEA PHARM INC.

XX
PT
DT
Wojcizn B. Ostretova-Goltz N. Lebowitz MS:

XX
TJ
B, 1120TOM
00721750

WPT: 2002-179695/23.

XX
PT
Determination of an agent capable of inhibiting aggregation of alpha
WEL, 2002-270033/42.

synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper -

XX Claim 40; Page 37; 52pp; English.

XX The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.

XX Sequence 7 AA;

Query Match 100.0%; Score 44; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 1 THIHPS 7

RESULT 2

ABP57617

ID ABP57617 standard; Protein; 148 AA.

AC ABP57617;

DT 29-APR-2003 (first entry)

DE S. murayamaensis ATCC 21414 kinamycin protein SEQ ID NO:19.

XX Streptomyces murayamaensis ATCC 21414; Streptomyces murayamaensis;
 KW glycosylated kinamycin; kinamycin; type II polyketide; polyketide;
 KW antibacterial; cytostatic; infection; antibiotic; antitumour;
 KW electrophilic azo-coupling agent.

XX Streptomyces murayamaensis.

PN WO2003002066-A2.

XX 09-JAN-2003.

XX 27-JUN-2002; 2002WO-US20719.

XX 27-JUN-2001; 2001US-301401P.

PA (DIVE-) DIVERSA CORP.

XX Short JM, Paradkar A, Varoglu M, Mathur EJ;

XX WPI; 2003-210195/20.

DR N-PSDB; ABZ71140.

XX New isolated polyketide used e.g. as antibiotic and antitumor agents
 PT comprises kinamycin molecule comprising at least one saccharide group

XX Claim 77; Page 92; 119pp; English.

XX The present invention describes a polyketide comprising a kinamycin
 CC molecule comprising at least one saccharide group. ABZ71132 to ABZ71163
 CC encode glycosylated kinamycins ABP57609 to ABP57640 isolated from
 CC Streptomyces murayamaensis ATCC 21414. Kinamycins are a class of
 CC type II polyketides. The kinamycins have antibacterial and cytostatic
 CC activities. They can be used for treating infections as antibiotics
 CC and as antitumour agents, and as electrophilic azo-coupling agents
 CC in vitro or in vivo.

XX Sequence 148 AA;

Query Match 97.7%; Score 43; DB 24; Length 148;

Best Local Similarity 85.7%; Pred. No. 2.4;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 93 THVHPS 99

RESULT 3

AAU19257

ID AAU19257 standard; Protein; 178 AA.

XX AAU19257;

XX 04-DEC-2001 (first entry)

DE Human G protein-coupled receptor nGPCR-2434.

XX Human; G protein-coupled receptor; nGPCR-x; antiviral; analgesic;
 KW cytotatic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive;
 KW antiparkinsonian; nootropic; neuroprotective; antidepressant;
 KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;
 KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;
 KW obesity; anorexia; hypotension; hypertension; myocardial infarction;
 KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;
 KW schizophrenia; migraine; major depression; anxiety; mental disorder;
 KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.

XX Homo sapiens.

XX WO200166750-A2.

XX 13-SEP-2001.

XX 08-MAR-2001; 2001WO-US07322.

XX 08-MAR-2000; 2000US-0187581.

XX 08-MAR-2000; 2000US-0187582.

XX 08-MAR-2000; 2000US-0187714.

XX 08-MAR-2000; 2000US-0187715.

XX 08-MAR-2000; 2000US-0187825.

XX 08-MAR-2000; 2000US-0187828.

XX 08-MAR-2000; 2000US-0187829.

XX 08-MAR-2000; 2000US-0187830.

XX 08-MAR-2000; 2000US-0187833.

XX 08-MAR-2000; 2000US-0187874.

XX 08-MAR-2000; 2000US-0187930.

XX 08-MAR-2000; 2000US-0188049.

XX 08-MAR-2000; 2000US-0189294.

XX 08-MAR-2000; 2000US-0187929.

XX 08-MAR-2000; 2000US-0187928.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS;

XX WPI; 2001-536779/59.

XX N-PSDB; AAS30826.

XX Isolated nucleic acid molecules encoding G protein-coupled receptors
 PT termed nGPCR-x, useful in the treatment and diagnosis of viral
 PT infections, cancers and mental disorders (e.g. Parkinson's disease and
 PT schizophrenia) -

XX Claim 31; Page 312-313; 336pp; English.

XX The invention relates to novel isolated nucleic acid molecules encoding
 CC G protein-coupled receptors termed nGPCR-x, nGPCR-x polynucleotides,
 CC polypeptides, and modulators may be used in the treatment of diseases and

CC conditions such as infections, such as viral infections caused by HIV-1
 CC (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and
 CC cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity,
 CC anorexia, hypotension, hypertension, myocardial infarction,
 CC atherosclerosis), Parkinson's disease, and psychotic and
 CC neurological disorders, including schizophrenia, migraine, major
 CC depression, anxiety, mental disorder, manic depression, and
 CC dyskinesias, such as Huntington's disease or Tourette's Syndrome
 CC and many other diseases and syndromes listed in the specification.
 CC nPCR-x polynucleotides and polypeptides, as well as nPCR-x
 CC modulators, may also be used in diagnostic assays for such diseases or
 CC conditions. The present sequence represents a G protein-coupled
 CC receptor of the invention.

XX Sequence 178 AA;

Query Match 88.6%; Score 39; DB 22; Length 178;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPS 7

Db 102 HHHPS 107

RESULT 4

ABG18105
 ID ABG18105 standard; Protein; 248 AA.

XX AC ABG18105;

XX DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #18096.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PS 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS82292.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 48464; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 248 AA;

Query Match 88.6%; Score 39; DB 22; Length 248;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPS 7

Db 61 HHHPS 66

RESULT 5

ABG19187

ID ABG19187 standard; Protein; 301 AA.

XX AC ABG19187;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #19178.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PS 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS83374.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 49546; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 301 AA;
 SQ Query Match 88.6%; Score 39; DB 22; Length 301;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 7
 || |||||
 Db 95 THCHPS 101

RESULT 6
 ABG21623
 ID ABG21623 standard; Protein; 301 AA.
 AC
 XX
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21614.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS85810.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 51982; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC a food supplement. (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 301 AA;
 SQ

Query Match 88.6%; Score 39; DB 22; Length 301;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 7
 || |||||
 Db 95 THCHPS 101

RESULT 7
 ABG21650
 ID ABG21650 standard; Protein; 322 AA.
 XX
 AC ABG21650;
 XX
 XX

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #21641.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX Homo sapiens.

XX WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS85837.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 20; SEQ ID No 52009; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC a food supplement. (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations in responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG00010-ABG30377 represent novel human CC diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 383 AA;
 Query Match 88.6%; Score 39; DB 22; Length 383;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 THIHPS 7
 |||||
 Db 247 THCHPS 253
 |||||
 RESULT 11
 AAR41310
 ID AAR41310 standard; Protein; 402 AA.
 XX AC AAR41310;
 XX DT 26-APR-1994 (first entry)
 XX DE Transcriptase.
 XX KW Retrotransposon; transcriptase; gene destruction; ss.
 XX OS Bacillus stearothermophilus.
 XX PN JP05227974-A.
 XX PD 07-SEP-1993.
 XX PF 24-FEB-1992; 92JP-0036426.
 XX PR 24-FEB-1992; 92JP-0036426.
 XX PA (MARU-) MARUKAN SU KK.
 XX DR WPI; 1993-316610/40.
 XX DR N-PSDB; AAQ49161.
 XX PT Retrotransposon contg., gene encoding transcriptase - has base
 sequence of C at position 10 to G at position 1339
 XX PS Claim 2; Page 4-6; 7pp; Japanese.
 XX CC The transcriptase gene, encoded by bases 118-1323 is claimed
 (claim 4) as is the retrotransposon containing bases 10-1339 of
 the sequence shown (claim 1). The transcriptase is useful for studying
 the mechanism of gene destruction or expression in Bacillus species.
 SQ Sequence 402 AA;
 Query Match 88.6%; Score 39; DB 14; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 THIHPS 7
 |||||
 Db 247 THCHPS 253
 |||||
 RESULT 12
 ABG02252
 ID ABG02252 standard; Protein; 437 AA.
 XX AC ABG02252;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #2243.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 383 AA;
 Query Match 88.6%; Score 39; DB 22; Length 383;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 THIHPS 7
 |||||
 Db 10 THCHPS 16
 |||||
 RESULT 10
 ABG19176
 ID ABG19176 standard; Protein; 383 AA.
 XX AC ABG19176;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #19167.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS83363.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity -
 Claim 20; SEQ ID No 49535; 103pp; English.
 The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful for treating
 a food supplement. (II) and (I) are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human
 diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences.

XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI N-PSDB; AAS66439.
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS66439.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 32611; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 437 AA;
Query Match 88.6%; Score 39; DB 22; Length 437;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 247 THCHPS 253
RESULT 13
ABG19205
ID ABG19205 standard; Protein; 451 AA.
XX AC ABG19205;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19196.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.

XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI N-PSDB; AAS83392.
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83392.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 49564; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 451 AA;
Query Match 88.6%; Score 39; DB 22; Length 451;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 247 THCHPS 253
RESULT 14
ABG10309
ID ABG10309 standard; Protein; 455 AA.
XX AC ABG10309;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #10300.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.

XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS95788.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 51960; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 478 AA;
Query Match 88.6%; Score 39; DB 22; Length 478;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THHHPS 7
Db 323 THCHPS 329
Search completed: February 11, 2004, 17:03:04
Job time : 33.25 secs

XX 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS74496.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 40668; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 455 AA;
Query Match 88.6%; Score 39; DB 22; Length 455;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THHHPS 7
Db 238 THCHPS 244
RESULT 15
ABG21601
ID ABG21601 standard; Protein; 478 AA.
XX AC ABG21601;
XX 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21592.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	39	88.6	646	4	US-09-625-188-10	US-09-625-188-10	Sequence 10, Appli
2	35	79.5	253	4	US-09-252-991A-20052	US-09-252-991A-20052	Sequence 20052, A
3	35	79.5	315	4	US-09-252-991A-19219	US-09-252-991A-19219	Sequence 19219, A
4	35	79.5	441	4	US-09-252-991A-32115	US-09-252-991A-32115	Sequence 32115, A
5	35	79.5	613	4	US-09-328-352-7191	US-09-328-352-7191	Sequence 7191, Ap
6	34	77.3	172	4	US-09-252-991A-22814	US-09-252-991A-22814	Sequence 22814, A
7	34	77.3	338	3	US-09-330-611-6	US-09-330-611-6	Sequence 6, Appli
8	34	77.3	344	2	US-08-755-728-3	US-08-755-728-3	Sequence 3, Appli
9	34	77.3	344	2	US-08-974-655-3	US-08-974-655-3	Sequence 3, Appli
10	34	77.3	344	3	US-08-283-011-3	US-08-283-011-3	Sequence 3, Appli
11	34	77.3	347	2	US-09-016-000-1	US-09-016-000-1	Sequence 1, Appli
12	34	77.3	511	4	US-09-252-991A-28223	US-09-252-991A-28223	Sequence 28223, A
13	34	77.3	558	4	US-09-252-991A-21305	US-09-252-991A-21305	Sequence 21305, A
14	34	77.3	568	4	US-09-252-991A-27039	US-09-252-991A-27039	Sequence 27039, A
15	33	75.0	212	4	US-09-328-352-7678	US-09-328-352-7678	Sequence 7678, Ap
16	33	75.0	255	4	US-09-031-962D-4	US-09-031-962D-4	Sequence 4, Appli
17	33	75.0	279	3	US-08-963-901-4	US-08-963-901-4	Sequence 4, Appli
18	33	75.0	296	4	US-09-252-991A-18620	US-09-252-991A-18620	Sequence 18620, A
19	33	75.0	299	4	US-09-252-991A-17615	US-09-252-991A-17615	Sequence 17615, A
20	33	75.0	315	4	US-09-252-991A-20012	US-09-252-991A-20012	Sequence 20012, A
21	33	75.0	329	4	US-09-252-991A-18007	US-09-252-991A-18007	Sequence 18007, A
22	33	75.0	361	4	US-09-252-991A-27989	US-09-252-991A-27989	Sequence 27989, A
23	33	75.0	391	4	US-09-134-001C-5234	US-09-134-001C-5234	Sequence 5234, Ap
24	33	75.0	393	4	US-09-252-991A-19043	US-09-252-991A-19043	Sequence 19043, A
25	33	75.0	410	4	US-09-252-991A-25451	US-09-252-991A-25451	Sequence 25451, A
26	33	75.0	457	3	US-08-963-901-2	US-08-963-901-2	Sequence 2, Appli
27	33	75.0	520	4	US-09-252-991A-24248	US-09-252-991A-24248	Sequence 24248, A

28 33 75.0 672 4 US-09-252-991A-24905 Sequence 24905, A

29 33 75.0 859 4 US-09-369-364A-5 Sequence 5, Appli

30 32 72.7 149 4 US-09-252-991A-19603 Sequence 19603, A

31 32 72.7 263 4 US-09-328-352-4326 Sequence 4326, Ap

32 32 72.7 449 2 US-08-927-394-2 Sequence 2, Appli

33 32 72.7 1315 4 US-09-252-991A-22746 Sequence 22746, A

34 32 72.7 3665 2 US-08-222-617A-13 Sequence 13, Appli

35 32 72.7 3712 2 US-08-222-617A-4 Sequence 4, Appli

36 32 72.7 3712 2 US-08-222-617A-25 Sequence 25, Appli

37 31 70.5 94 4 US-09-252-991A-31551 Sequence 31551, A

38 31 70.5 141 4 US-09-353-719-2 Sequence 2, Appli

39 31 70.5 217 4 US-09-252-991A-24723 Sequence 24723, A

40 31 70.5 292 4 US-08-858-207A-391 Sequence 391, App

41 31 70.5 357 4 US-08-252-991A-22404 Sequence 22404, A

42 31 70.5 456 4 US-09-328-352-4944 Sequence 4944, Ap

43 31 70.5 538 4 US-09-252-991A-20413 Sequence 20413, A

44 31 70.5 565 4 US-09-506-286B-8 Sequence 8, Appli

45 31 70.5 565 4 US-09-506-286B-11 Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-625-188-10
; Sequence 10, Application US/09625188
; Patent No. 6307037
; GENERAL INFORMATION:
; APPLICANT: No. 6307037artis AG
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-31285P1
; CURRENT APPLICATION NUMBER: US/09/625.188
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Ashbya gossypii
US-09-625-188-10

Query Match 88.6%; Score 39; DB 4; Length 646;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7

DB 304 HIHPS 309

RESULT 2

US-09-252-991A-20052
; Sequence 20052, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33.42
; SEQ ID NO 20052
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20052

Query Match 79.5%; Score 35; DB 4; Length 253;

Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6
|||
Db 26 HHHP 30

RESULT 3
US-09-252-991A-19219
; Sequence 19219, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19219
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19219

Query Match 79.5%; Score 35; DB 4; Length 315;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 7
|||
Db 307 HVHHP 312

RESULT 4
US-09-252-991A-32115
; Sequence 32115, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32115
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32115

Query Match 79.5%; Score 35; DB 4; Length 441;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 7
|||
Db 277 HVHHP 282

RESULT 5
US-09-328-352-7191
; Sequence 7191, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7191
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7191

Query Match 79.5%; Score 35; DB 4; Length 613;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHP 6
|||
Db 579 THHHP 584

RESULT 6
US-09-252-991A-22814
; Sequence 22814, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22814
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22814

Query Match 77.3%; Score 34; DB 4; Length 172;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6
|||
Db 2 HVHHP 6

RESULT 7
US-09-330-611-6
; Sequence 6, Application US/09330611
; Patent No. 6248874
; GENERAL INFORMATION:
; APPLICANT: FREY, Perry A.
; APPLICANT: RUZICKA, Frank J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING BACTERIAL LYSINE 2,3-AMINOMUTASE
; FILE REFERENCE: 032026/0476
; CURRENT APPLICATION NUMBER: US/09/330,611
; CURRENT FILING DATE: 1999-06-11
; EARLIER APPLICATION NUMBER: US 09/198,942
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-330-611-6

Query Match 77.3%; Score 34; DB 3; Length 338;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THIHPS 7
Db 225 THINHPN 231

RESULT 8

US-08-755-728-3
; Sequence 3, Application US/08755728

; Patent No. 5962312

; GENERAL INFORMATION:

; APPLICANT: Plowman, Gregory

; APPLICANT: Mossie, Kevin

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1

; AND/OR AUR-2 RELATED DISORDERS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: California

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSEQ for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/755.728

; FILING DATE: No. 5962312ember 25, 1996

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/008.809

; FILING DATE: December 18, 1995

; APPLICATION NUMBER: 60/023.943

; FILING DATE: August 14, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 223/113

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 344 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-755-728-3

Query Match 77.3%; Score 34; DB 2; Length 344;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHPS 7

Db 131 HLHHPN 136

RESULT 9

US-08-974-655-3

; Sequence 3, Application US/08974655

; Patent No. 5972676

; GENERAL INFORMATION:

; APPLICANT: Plowman, Gregory

; APPLICANT: Mossie, Kevin

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1

; AND/OR AUR-2 RELATED DISORDERS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: California

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSEQ for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/974.655

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/755.728

; FILING DATE: No. 5972676ember 25, 1996

; APPLICATION NUMBER: 60/008.809

; FILING DATE: December 18, 1995

; APPLICATION NUMBER: 60/023.943

; FILING DATE: August 14, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 223/113

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 344 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-974-655-3

Query Match 77.3%; Score 34; DB 2; Length 344;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHPS 7

Db 131 HLHHPN 136

RESULT 10

US-09-283-011-3

; Sequence 3, Application US/09283011

; Patent No. 6207401

; GENERAL INFORMATION:

; APPLICANT: Plowman, Gregory

APPLICANT: Mossier, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
SEQUENCE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/283,011
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/012,135
FILING DATE: January 22, 1998
APPLICATION NUMBER: 08/755,728
FILING DATE: No 6207401ember 25, 1996
APPLICATION NUMBER: 60/023,943
FILING DATE: August 14, 1996
APPLICATION NUMBER: 60/008,809
FILING DATE: December 18, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 211/282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-283-011-3

Query Match 77.3%; Score 34; DB 3; Length 344;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPHS 7
Db 131 HLHFN 136

RESULT 11
US-09-016-000-1
Sequence 1, Application US/09016000
Patent No. 5962232
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
TITLE OF INVENTION: PROTEIN KINASE MOLECULES
NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,000
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PE-0465 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMCINOT01
CLONE: 2940
US-09-016-000-1

Query Match 77.3%; Score 34; DB 2; Length 347;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPHS 7
Db 131 HLHFN 136

RESULT 12
US-09-252-991A-28223
Sequence 28223, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28223
LENGTH: 511
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28223

Query Match 77.3%; Score 34; DB 4; Length 511;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|:|
Db 318 HVHHP 322

RESULT 13

US-09-252-991A-21305
; Sequence 21305, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21305

; LENGTH: 558

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21305

Query Match 77.3%; Score 34; DB 4; Length 558;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 7
|:|
Db 345 HVHHP 350

RESULT 14

US-09-252-991A-27039

; Sequence 27039, Application US/09252991A

; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27039

; LENGTH: 568

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27039

Query Match 77.3%; Score 34; DB 4; Length 568;

Best Local Similarity 80.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|:|
Db 439 HVHHP 443

RESULT 15

US-09-328-352-7678

; Sequence 7678, Application US/09328352

; Patent No. 6562958

GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7678
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7678

Query Match 75.0%; Score 33; DB 4; Length 212;

Best Local Similarity 80.0%; Pred. No. 1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|:|
Db 100 HLHHP 104

Search completed: February 11, 2004, 17:13:39

Job time : 10.4167 secs